

# LOGARITHMIC $L^p$ BOUNDS FOR MAXIMAL DIRECTIONAL SINGULAR INTEGRALS IN THE PLANE

CIPRIAN DEMETER AND FRANCESCO DI PLINIO

ABSTRACT. Let  $K$  be a Calderon-Zygmund convolution kernel on  $\mathbb{R}$ . We discuss the  $L^p$ -boundedness of the maximal directional singular integral

$$T_{\mathbf{V}}f(x) = \sup_{v \in \mathbf{V}} \left| \int_{\mathbb{R}} f(x + tv)K(t) dt \right|$$

where  $\mathbf{V}$  is a finite set of  $N$  directions. Logarithmic bounds (for  $2 \leq p < \infty$ ) are established for a set  $\mathbf{V}$  of arbitrary structure. Sharp bounds are proved for lacunary and Vargas sets of directions. The latter include the case of uniformly distributed directions and the finite truncations of the Cantor set.

We make use of both classical harmonic analysis methods and product-BMO based time-frequency analysis techniques. As a further application of the latter, we derive an  $L^p$  almost orthogonality principle for Fourier restrictions to cones.

## 1. INTRODUCTION AND MAIN RESULTS

**1.1. Maximal directional singular integrals.** Let  $m$  be a Hörmander-Mikhlin multiplier on  $\mathbb{R}$ . For  $f \in \mathcal{C}_0^\infty(\mathbb{R}^2)$ , and  $v \in S^1$ , consider the directional multiplier

$$(1.1) \quad T_v f(x_1, x_2) = \int_{\mathbb{R}^2} \hat{f}(\xi_1, \xi_2) m(\xi \cdot v) e^{i(x_1 \xi_1 + x_2 \xi_2)} d\xi_1 d\xi_2.$$

The multiplier  $T_v$  admits the singular integral representation

$$T_v f(x) = \int_{\mathbb{R}} f(x + tv) K(t) dt, \quad x \notin \text{supp } f,$$

where  $K = \check{m}$  is a Calderon-Zygmund kernel. The most basic example is the directional Hilbert transform

$$H_v f(x) = \int_{\mathbb{R}} f(x + tv) \frac{dt}{t}, \quad x \notin \text{supp } f,$$

which corresponds to the multiplier  $\text{sign}(\xi)$ . Assume now that we have a vector field  $\mathbf{v} : \mathbb{R}^2 \rightarrow S^1$  and define the corresponding directional singular integral operator

$$T_{\mathbf{v}}(f)(x) = T_{\mathbf{v}(x)}f(x).$$

Various cases where the vector field satisfies certain geometric or analytic constraints have been recently investigated. In [23] and [20], Lacey and Li developed a beautiful theory for

---

1991 *Mathematics Subject Classification*. Primary: 42B20; Secondary: 42B25.

*Key words and phrases*. Maximal singular integrals, Keakeya maximal function.

The first author is partially supported by a Sloan Research Fellowship and by NSF Grant DMS-0901208. The second author was partially supported by the National Science Foundation under the grant NSF-DMS-0906440, and by the Research Fund of Indiana University.

$C^{1+\epsilon}$  vector fields. Their analysis revealed intricate connections between the directional Hilbert transform, Carleson's theorem on the convergence of Fourier series and a certain restricted version of the Kakeya maximal function.

The Lacey-Li techniques have been recently refined by Bateman and Thiele. They combined novel density estimates for rectangles with a new approach to proving vector valued inequalities, to obtain  $L^p$  estimates for  $H_{\mathbf{V}}$  in the case when the vector field depends on just one coordinate [2]. On the other hand, Stein and Street [30] have obtained analogous results for the case of real analytic vector fields.

Continuing the initial investigation from [6], in this paper we analyze another case of interest, the one where the range of the vector field is finite. More precisely, we are concerned with  $L^p$  bounds for the maximal directional multiplier

$$(1.2) \quad T_{\mathbf{V}}f(x) = \sup_{v \in \mathbf{V}} |T_v f(x)|,$$

where  $\mathbf{V} \subset S^1$  is a set of  $N$  directions. Due to the special role played by the maximal directional Hilbert transform, we will reserve for it the notation  $H_{\mathbf{V}}$ .

It has been long recognized that there is a close connection between maximal and singular integral operators in harmonic analysis, and indeed, in most classical cases the two types of operators have similar boundedness properties. In many ways this paper will argue against the generality of this principle. We start by briefly recalling some history. The relevant directional maximal function

$$(1.3) \quad M_{\mathbf{V}}f(x) = \sup_{v \in \mathbf{V}} |M_v f(x)| \quad M_v f(x) = \sup_{\varepsilon > 0} \frac{1}{2\varepsilon} \left| \int_{-\varepsilon}^{\varepsilon} f(x + tv) dt \right|,$$

has been intensely studied and is now completely understood. The sharp bounds (for generic  $\mathbf{V}$ )

$$(1.4) \quad \|M_{\mathbf{V}}\|_{2 \rightarrow 2} \lesssim \log N, \quad \|M_{\mathbf{V}}\|_{2 \rightarrow 2, \infty} \lesssim \sqrt{\log N}, \quad \|M_{\mathbf{V}}\|_{p \rightarrow p} \lesssim (\log N)^{\frac{1}{p}}, \quad p > 2,$$

were first proved in [17]. The more recent work [1] shows that bounds become independent of  $N$

$$\|M_{\mathbf{V}}\|_{p \rightarrow p} \lesssim_p 1, \quad 1 < p < \infty,$$

if and only if  $\mathbf{V}$  is lacunary of finite order.

In striking contrast to  $M_{\mathbf{V}}$ , the sharp (in terms of  $N$ ) upper bounds for the operator norms  $\|H_{\mathbf{V}}\|_{p \rightarrow p}$  are only known for  $1 < p \leq 2$ . Taking  $f$  to be the indicator function of the unit ball and  $\mathbf{V}$  uniformly distributed, one gets the lower bounds (see [6])

$$\|H_{\mathbf{V}}f\|_p \gtrsim N^{\frac{1}{p}} \|f\|_p, \quad 1 < p < 2, \quad \|H_{\mathbf{V}}f\|_2 \gtrsim \log N \|f\|_2.$$

These lower bounds are in fact also upper bounds for generic  $\mathbf{V}$ . The case  $p = 2$  is a simple consequence of the Rademacher-Menshov theorem. Indeed,  $|H_{\mathbf{V}}f(x)|$  is bounded by the sum of two operators of the form

$$(1.5) \quad \max_{1 \leq \nu \leq N} \left| \sum_{j=1}^{\nu} f_j(x) \right|$$

where  $f_j$  are Fourier restrictions of  $f$  to disjoint frequency cones rooted at the origin. On the other hand, for each  $N$  there exists an orthonormal set  $\{f_j : j = 1, \dots, N\}$  with

$$\left| \left\{ x \in \mathbb{R}^2 : \max_{1 \leq \nu \leq N} \left| \sum_{j=1}^{\nu} f_j(x) \right| \gtrsim \sqrt{N} \log N \right\} \right| > 1$$

(this is Menshov's counterexample, see [16]); thus one cannot get a weak  $L^2$  bound for  $H_{\mathbf{V}}$  which is better than the strong  $L^2$  bound by simply invoking Rademacher-Menshov type results.

In the general case, the sharp bound

$$\|T_{\mathbf{V}}f\|_2 \lesssim \log N \|f\|_2$$

has been established in [6]. Quite surprisingly, the article [15] proves the lower bound

$$(1.6) \quad \|H_{\mathbf{V}}\|_{2 \rightarrow 2, \infty} \gtrsim \sqrt{\log N}, \quad \forall \mathbf{V} \subset S^1, \# \mathbf{V} = N.$$

Thus  $H_{\mathbf{V}}$  is an unbounded operator on  $L^2$  as soon as  $\mathbf{V}$  contains infinitely many directions, in particular for infinite lacunary sets. This is in sharp contrast with the directional maximal function.

Our main goal is to investigate the sharp dependence on  $N$  of the  $L^p$  ( $2 < p < \infty$ ) and also weak  $L^2$  operator norms of  $T_{\mathbf{V}}$ . In analogy with the case of the maximal function (1.4), one expects logarithmic-type bounds for all  $2 \leq p < \infty$ . However, at least three structural differences between  $T_{\mathbf{V}}$  and  $M_{\mathbf{V}}$  make the former harder to treat. The maximal directional singular integral is not a positive operator; it is a *sum* not a *maximum* of single scale operators. Moreover, unlike in the case of  $M_{\mathbf{V}}$  (which is trivially bounded on  $L^\infty$  for any  $\mathbf{V}$ ), there is no immediate evidence of an endpoint result for  $T_{\mathbf{V}}$ .

We briefly mention some of the preexisting literature on maximal directional singular integrals. In [4], the authors deal with the boundedness of  $\{H_v f(x) : x \in \mathbb{R}^n, v \in S^{n-1}\}$  and of the relative maximal truncations in the mixed norm spaces  $L^p(\mathbb{R}^n; L^q(S^{n-1}))$ ,  $1 < p, q < \infty$ . In the article [9], it is proven that, whenever  $\mathbf{V} \subset S^1$  is a set with Minkowski dimension  $d(\mathbf{V}) < 1$ ,  $M_{\mathbf{V}}$  and  $H_{\mathbf{V}}$  are bounded operators on the subspace of the radial functions in  $L^p(\mathbb{R}^2)$  for  $p > 1 + d(\mathbf{V})$ , and unbounded for  $p < 1 + d(\mathbf{V})$ . The more recent article [19] reproves that  $\|H_{\mathbf{V}}\|_{2 \rightarrow 2} \sim \log N$  when  $\# \mathbf{V} = N$  without appealing to Rademacher-Menshov explicitly, and by the same method obtains the sharp  $L^2$  bound (which is  $O(N^{\frac{n}{2}-1})$ ) for the analogous maximal directional Hilbert transform in  $\mathbb{R}^n$ ,  $n > 2$ . Our results partially answer both open problems mentioned in [19].

**1.2. Notation.** Throughout the paper,  $m$  stands for a Hörmander-Mikhlin multiplier on  $\mathbb{R}$ , that is

$$m \in C^\infty(\mathbb{R} \setminus \{0\}), \quad |\partial^\alpha m(\xi)| \lesssim_\alpha |\xi|^{-\alpha}, \quad \alpha \geq 0;$$

with corresponding multiplier operator

$$T : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}), \quad Tf(x) = \int_{\mathbb{R}} \hat{f}(\xi) m(\xi) e^{ix\xi} d\xi.$$

We will also make use of the maximal truncation of  $T$ , and adopt the notation

$$T_\star f(x) = \sup_{\varepsilon > 0} |T_\varepsilon f(x)|, \quad T_\varepsilon f(x) = \int_{|t| > \varepsilon} f(x-t) K(t) dt, \quad K = \check{m}.$$

The values of the positive constants  $C, c$ , as well as the implicit constants hidden in the notation  $\lesssim$  may vary from line to line, and, unless otherwise specified, are allowed to depend only on the multiplier  $m$ .

The Hardy-Littlewood maximal function of  $f$  on  $\mathbb{R}^n$  ( $n = 1, 2$ ) is denoted by  $Mf$ . For  $R \subset \mathbb{R}^1$  or  $\mathbb{R}^2$ , we set  $\mathbb{E}_R f = \frac{1}{|R|} \int_R f$ , and, for a collection  $\mathcal{R}$  of rectangles in the plane, we denote the corresponding maximal function by

$$\mathbf{M}_{\mathcal{R}} f(x) = \sup_{R \in \mathcal{R}} |\mathbb{E}_R f| \mathbf{1}_R(x), \quad x \in \mathbb{R}^2.$$

**1.3. The main results.** The first main result concerns arbitrary finite sets of directions. We believe that the exponent one of the term  $\log N$  gets sharp in the limiting case  $p \rightarrow \infty$ , see Conjecture 1.3.

**Theorem 1.** *Let  $m$  be a Hörmander-Mihlin multiplier and let  $T_{\mathbf{V}}$  be defined as in (1.2). For any given set  $\mathbf{V}$  of  $N$  directions*

$$(1.7) \quad \|T_{\mathbf{V}}\|_{p \rightarrow p} \leq Cp \log N, \quad 2 < p < \infty.$$

*Moreover the following endpoint result holds: if  $\text{supp } f \subset Q \subset \mathbb{R}^2$*

$$(1.8) \quad |\{x \in Q : |T_{\mathbf{V}} f(x)| > \lambda \log N\}| \leq C \exp\left(-c \frac{\lambda}{\|f\|_{\infty}}\right) |Q|, \quad \forall \lambda > 0.$$

**Remark 1.1.** A simpler proof of this result for  $T_{\mathbf{V}} = H_{\mathbf{V}}$  is given in paragraph 6.1.3, exploiting the almost  $L^p$ -orthogonality principle in Theorem 4.

**Definition 1.2.** An ordered set  $\mathbf{V} = \{v_1, v_2, \dots\} \subset S^1$  (finite or countably infinite) is called *lacunary* with node  $v_{\infty} \in S^1$  if

$$(1.9) \quad |v_{j+1} - v_{\infty}| \leq \frac{1}{2} |v_j - v_{\infty}|, \quad j = 1, 2, \dots$$

**Definition 1.3.** [18] A finite set  $\mathbf{V} \subset S^1$  with  $\#\mathbf{V} = N$  is a *Vargas set* with constant  $Q$  if

$$(1.10) \quad \max\{\#\mathbf{V}' : \mathbf{V}' \text{ is a lacunary subset of } \mathbf{V}\} = Q \log N.$$

The Vargas sets with constants  $Q = O(1)$  independent of  $N$  will simply be referred to as Vargas sets. Examples include uniformly distributed sets of directions  $\mathbf{V}_N = \{v_j = e^{2\pi i \frac{j}{N}}, j = 0, \dots, N-1\}$  which are easily seen to be Vargas sets with  $Q = 1$ . The  $N$ -truncations of Cantor sets are also Vargas sets with  $Q = 4$ . See Subsection 7.1 for more details.

Finite lacunary sets and Vargas sets are at the opposite ends of the spectrum in terms of maximum cardinality of lacunary subsets. In fact, every set of  $N$  directions is “at least  $\frac{1}{3}$  Vargas”, in the sense that it contains a lacunary subsets with  $\frac{\log N}{3}$  elements (a proof is given in Subsection 7.1). In these two extreme (but rather relevant) cases, we are able to improve the result of Theorem 1 and obtain essentially sharp bounds.

**Theorem 2.** *Let  $\mathbf{V}$  be a lacunary set of  $N$  directions. We have the sharp bound*

$$(1.11) \quad \|T_{\mathbf{V}}\|_{p \rightarrow p} \lesssim_p \sqrt{\log N}, \quad 1 < p < \infty.$$

**Theorem 3.** *Let  $\mathbf{V}$ ,  $\#\mathbf{V} = N$ , be a Vargas set with constant  $Q$ . We have the following bound*

$$(1.12) \quad \|H_{\mathbf{V}}\|_{2 \rightarrow 2, \infty} \lesssim_Q \sqrt{\log N} (\log \log N)^6.$$

*In particular, if  $V$  is uniform or a Cantor set then the following essentially sharp bound holds*

$$\|H_{\mathbf{V}}\|_{2 \rightarrow 2, \infty} \lesssim \sqrt{\log N} (\log \log N)^6.$$

**Remark 1.4.** The (essential) sharpness of Theorem 3 follows immediately via comparison with (1.6). It is not clear whether the  $\log \log N$  term can be eliminated.

Note also that the bound in Theorem 3 loses its strength as  $Q$  gets large.

Interestingly, Theorem 2 and Theorem 3 show that uniformly distributed sets and lacunary sets of  $N$  directions have the same quantitative behavior on weak  $L^2$  (this is not the case for strong  $L^2$ ).

The methods of proof of Theorem 3 can also be easily modified to prove the estimate

$$\|H_{\mathbf{V}} f\|_p \lesssim_p (\log N)^{\frac{1}{p'}} (\log \log N)^{C_p}, \quad 2 < p < \infty,$$

for each Vargas set of cardinality  $N$ . We strongly believe these methods are sharp, and thus that the exponent of  $\log N$  is the correct one for each  $p > 2$ . We are thus led to the following conjecture, which, as our results show, holds for lacunary sets and (at least in the case of the Hilbert transform) Vargas sets.

**Conjecture.** *Let  $T_{\mathbf{V}}$  be defined as in (1.2) and  $\#\mathbf{V} = N$ . We have the bounds*

$$\|T_{\mathbf{V}}\|_{2 \rightarrow 2, \infty} \lesssim \sqrt{\log N} (\log \log N)^C, \quad \|T_{\mathbf{V}}\|_{p \rightarrow p} \lesssim_p (\log N)^{\frac{1}{p'}} (\log \log N)^{C_p}, \quad 2 < p < \infty.$$

In light of our Theorem 1, the second conjectured inequality will follow if the first inequality is established.

**1.4. The methods and the plan of the paper.** Theorem 1 and 2 are proved by classical harmonic analysis methods. The main tool behind the proof of Theorem 1, which is given in Section 2, is an exponential-type good-lambda inequality for one-dimensional singular integrals by Hunt [13], which allows to estimate the contribution of the  $N$  singular integrals  $T_v$  with a loss which is only logarithmic.

Theorem 2 is proved in Section 3. Via an application of the exponential square integrability inequality by Chang, Wilson, and Wolff [3], we lose a  $\sqrt{\log N}$  and reduce  $T_{\mathbf{V}}$  to a square function in the frequency Littlewood-Paley annuli, which is shown to be uniformly bounded in  $N$  when  $\mathbf{V}$  is lacunary.

On the other hand, Theorem 3 is proved by means of time-frequency analysis in the phase plane. We use a discretization for  $H_{\mathbf{V}}$  very similar to the one introduced by Lacey and Li in their pioneering work [23] and [20]. The resulting model sum closely resembles the one introduced in [22] for Carleson's operator (see Sections 5 and 6). However, the *tiles* appearing in the model sum for  $H_{\mathbf{V}}$  are a bi-parameter family, and are allowed to be oriented along each direction in  $\mathbf{V}$ . These model sums are chopped into *single tree* operators. Our analysis is somewhat similar to the one from [8] which contains a new proof of the boundedness of the Carleson operator. The approach from [8] is to show that for each point  $x$ , the contribution to the Carleson operator comes essentially from one

tree. For Vargas sets, there will be at most  $\log N$  trees contributing to  $H_{\mathbf{V}}$ . This upper bound is further lowered to  $\sqrt{\log N}$  by interpolating with a Bessel inequality.

One of the things that distinguishes our approach from the ones in [2], [6], [23] and [20] is the fact that our trees will be of product nature (i.e. bi-parameter). The bulk of the argument is given in Section 7.

## 2. PROOF OF THEOREM 1

Throughout this section, we use the notations

$$\mu_Q(A) = \int_Q \mathbf{1}_A(x) \frac{dx}{|Q|}, \quad \|f\|_{L^p(Q)} := \left( \int_Q |f(x)|^p d\mu_Q(x) \right)^{\frac{1}{p}} = \left( \int_Q |f(x)|^p \frac{dx}{|Q|} \right)^{\frac{1}{p}}$$

for  $A \subset \mathbb{R}^2$  and  $1 \leq p < \infty$ . Before we enter the proof, we recall two results we cited in the introduction, which will be used below: the sharp bound

$$(2.1) \quad \|T_{\mathbf{V}}f\|_{2 \rightarrow 2} \lesssim \log N$$

which has been proved in [6], and

$$(2.2) \quad \|M_{\mathbf{V}}f\|_{p \rightarrow p} \leq C(\log N)^{\frac{1}{p}} \|f\|_p, \quad 2 < p < \infty$$

which has first been established in [17].

**2.1. Maximal truncations of CZ kernels and a good- $\lambda$  inequality.** To prove Theorem 1, we need two propositions. The first is a consequence of the sharp weighted bound for maximal truncations of singular integrals in terms of the weight characteristic. The second is a reformulation of a classical good-lambda inequality by Hunt [13].

**Proposition 2.1.** *We have the bound  $\|T_{\star}f\|_p \leq Cp\|f\|_p$ , for each  $2 < p < \infty$ .*

**Proposition 2.2.** *There exist positive absolute constants  $c_0, c_1$  such that for all  $f \in \mathcal{C}_0^\infty(\mathbb{R})$ , all  $\lambda > 0$  and all  $\gamma > 4$ ,*

$$(2.3) \quad |\{x \in \mathbb{R} : |Tf(x)| > \gamma\lambda, Mf(x) < c_0\lambda\}| \leq Ce^{-c_1\gamma} |\{x \in \mathbb{R} : |T_{\star}f(x)| > \lambda\}|.$$

*Proof of Proposition 2.1.* As in [10], we follow Rubio de Francia's method. The recent article [14] contains the sharp bound

$$(2.4) \quad \|T_{\star}f\|_{L^2(w)} \leq C[w]_{A_2} \|f\|_{L^2(w)},$$

where  $[w]_{A_2}$  denotes the  $A_2$  characteristic of the weight  $w$ . Let  $q = (p/2)' > 1$ . Choose  $v \in L^q(\mathbb{R})$ , such that  $\|v\|_q = 1$  and

$$\|T_{\star}f\|_p^2 = \int_{\mathbb{R}} |T_{\star}f(x)|^2 v(x) dx.$$

By interpolating the trivial  $L^\infty$  bound with the endpoint  $\|Mf\|_1 \leq C\|f\|_{L \log L}$ , one obtains the (sharp) bound  $\|M\|_{q \rightarrow q} \leq Cp$ . Therefore, defining

$$w = \sum_{k=0}^{\infty} \frac{1}{(2\|M\|_{q \rightarrow q})^k} \underbrace{M \circ \dots \circ M}_{k \text{ times}} v$$

we have

$$v(x) \leq w(x) \text{ a.e., } \|w\|_q \leq 2, \quad Mw(x) \leq Cp w(x) \text{ a.e., } [w]_{A_2} \leq Cp.$$

We can conclude the proof with the chain of inequalities

$$\begin{aligned} \|T_\star f\|_p^2 &= \int_R |T_\star f(x)|^2 v(x) dx \leq \int_R |T_\star f(x)|^2 w(x) dx \\ &\leq C^2 [w]_{A_2}^2 \int_R |f(x)|^2 w(x) dx \\ &\leq C^2 p^2 \|f\|_p^2, \end{aligned}$$

where we used (2.4) in going from the first to the second line.  $\square$

*Proof of Proposition 2.2.* Let us call  $G$  the set on the right-hand side of (2.3). We can write  $G$  as the countable union of intervals  $J$  with disjoint interiors such that  $3J \not\subset G$ . Therefore, (2.3) will follow by summing up the estimate

$$(2.5) \quad |\{x \in J : |Tf(x)| > \gamma\lambda, Mf(x) < c_0\lambda\}| \leq Ce^{-c_1\gamma}|J|.$$

Let us prove (2.5). Choose a point  $\bar{x} \in 3J \cap G^c$  and let  $\bar{J}$  be the interval of length  $6|J|$  centered at  $\bar{x}$ . Set

$$f = f_1 + f_2, \quad f_1 = f\mathbf{1}_{\bar{J}}, \quad f_2 = f\mathbf{1}_{(\bar{J})^c}.$$

It is clear that

$$(2.6) \quad (2.5)_{\text{LHS}} \leq |\{x \in \bar{J} : |Tf_1(x)| > \tfrac{\gamma}{2}\lambda, Mf(x) < c_0\lambda\}| + |\{x \in J : |Tf_2(x)| > \tfrac{\gamma}{2}\lambda\}|$$

Rescaling Proposition 2 of [13] on  $\bar{J}$ , we get that

$$(2.7) \quad |\{x \in \bar{J} : |Tf_1(x)| > \tfrac{\gamma}{2}\lambda, Mf(x) < c_0\lambda\}| \leq Ae^{-\frac{a_0}{2c_0}\gamma}|\bar{J}| \leq 6Ae^{-\frac{a_0}{2c_0}\gamma}|J|.$$

for some positive absolute constants  $A, a_0$ . If we prove that the set in the second term of (2.6)<sub>RHS</sub> is empty for some choice of  $c_0$ , (2.5) will follow from (2.7), with  $C = 6A$ ,  $c_1 = \frac{a_0}{2c_0}$ . This is done as follows. It is known that (see [29], p. 208, eq. (35))

$$(2.8) \quad |Tf_2(\bar{x}) - Tf_2(x)| \leq A_1 \inf_J Mf \leq A_1 c_0 \lambda, \quad \forall x \in J,$$

for some absolute constant  $A_1 > 0$ . Here the second inequality follows because we can work under the assumption that the set in (2.5)<sub>LHS</sub> is nonempty. Also note that

$$Tf_2(\bar{x}) = T_\varepsilon f(\bar{x}), \quad \varepsilon = 6|J|$$

and by our choice of  $\bar{x} \notin G$ ,  $|T_\varepsilon f(\bar{x})| \leq T_\star f(\bar{x}) \leq \lambda$ . Hence, if we set  $c_0 = \frac{1}{A_1}$ , we get

$$|Tf_2(x)| \leq |Tf_2(\bar{x}) - Tf_2(x)| + T_\star f(\bar{x}) \leq 2\lambda \leq \frac{\gamma}{2}\lambda, \quad \forall x \in J,$$

and this completes the proof.  $\square$



**2.2. Proof of Theorem 1.** By a limiting argument, we can assume that  $f$  is supported on a square  $Q \subset \mathbb{R}^2$  of (finite) sidelength  $\ell(Q)$ . We first treat the part of  $T_{\mathbf{V}}f$  outside  $2Q$ . Let  $x \in 2^{k+1}Q \setminus 2^kQ$ ,  $k \geq 1$ . We have, using the decay properties of the kernel  $K$ ,

$$(2.9) \quad |T_v f(x)| \leq \int_{\delta_1(x)}^{\delta_2(x)} |f(x+tv)| |K(t)| dt \leq \frac{C}{2^k \ell(Q)} \int_{-2^k \ell(Q)}^{2^k \ell(Q)} |f(x+tv)| dt \leq C M_v |f|(x).$$

Here we assumed that the line  $\{x+tv : t \in \mathbb{R}\}$  intersects  $Q$  at all  $t$  between  $\delta_1(x)$  and  $\delta_2(x)$ , with  $\delta_1(x), \delta_2(x) \sim 2^k \ell(Q)$  (otherwise (2.9) holds trivially). Therefore

$$(2.10) \quad \|(T_{\mathbf{V}}f) \mathbf{1}_{\mathbb{R}^2 \setminus 2Q}\|_p \leq C \|M_{\mathbf{V}}|f|\|_p \leq C (\log N)^{\frac{1}{p}} \|f\|_p.$$

We now show that

$$(2.11) \quad \|T_{\mathbf{V}}f\|_{L^p(2Q)} \leq C p \log N \|f\|_{L^p(Q)}, \quad \forall 2 < p < \infty;$$

this and (2.10) immediately imply (1.7). The weak-type endpoint result (1.8) is then obtained as an easy consequence of (2.11) (proceeding as in, for instance, [29, Sect. IV.1.3]).

In short, to prove (2.11), we apply the exponential good-lambda inequality of Proposition 2.2 for each  $T_v$ , which acts separately on each fiber of  $f$  along  $v$ , to excise  $N$  different exceptional sets. Due to the exponential decay in the good-lambda inequality, the final loss is only logarithmic.

*Proof of (2.11).* The range of exponents is open, so by Marcinkiewicz interpolation it suffices to prove the weak-type bound

$$(2.12) \quad \mu_{2Q}(\{x \in 2Q : T_{\mathbf{V}}f(x) > \lambda \log N\}) \leq \frac{C^p p^p}{\lambda^p} \|f\|_{L^p(Q)}^p, \quad \lambda > 0, 2 < p < \infty.$$

For  $\lambda > 0$ , split ( $c_0$  below is the absolute constant appearing in Proposition 2.2)

$$f = f_1 + f_2, \quad f_1(x) = f(x) \mathbf{1}_{\{|f| \leq c_0 \lambda\}}, \quad f_2 = f - f_1,$$

We get at once that

$$(2.13) \quad \sup_{x \in Q} |f_1(x)| \leq c_0 \lambda,$$

$$(2.14) \quad \|f_1\|_{L^p(Q)} \leq \|f\|_{L^p(Q)},$$

$$(2.15) \quad \|f_2\|_{L^2(Q)}^2 \leq C \lambda^{2-p} \|f\|_{L^p(Q)}^p,$$

$$(2.16) \quad \mu_{2Q}(\{T_{\mathbf{V}}f > 2\lambda \log N\}) \leq \mu_{2Q}(\{T_{\mathbf{V}}f_1 > \lambda \log N\}) + \mu_{2Q}(\{T_{\mathbf{V}}f_2 > \lambda \log N\})$$

We estimate the second term in the right-hand side of (2.16); using the  $L^2$  bound of (2.1), and (2.15),

$$(2.17) \quad \mu_{2Q}(\{T_{\mathbf{V}}f_2 > \lambda \log N\}) \leq \frac{C}{\lambda^2} \|f_2\|_{L^2(Q)}^2 \leq \frac{C}{\lambda^p} \|f\|_{L^p(Q)}^p.$$

We turn to the first term in the right-hand side of (2.16). Preliminarily observe that

$$\mu_{2Q}(\{T_{\mathbf{V}}f_1 > \lambda \log N\}) \leq \mu_{2Q}(\{T_{\mathbf{V}}f_1 > \lambda \log N, M_{\mathbf{V}}f < c_0 \lambda\}) + \mu_{2Q}(\{M_{\mathbf{V}}f > c_0 \lambda\}),$$



and (2.13) implies that the second set on the right hand side is empty. Thus, in view of (2.17), (2.12) follows from the estimate

$$(2.18) \quad \mu_{2Q}(\{x \in \mathbb{R}^2 : T_{\mathbf{V}}g(x) > \lambda \log N, M_{\mathbf{V}}g(x) < c_0\lambda\}) \leq \frac{C^p p^p}{\lambda^p} \|g\|_{L^p(Q)}^p,$$

applied to  $g = f_1$ . We have completed the proof of (2.12) (and thus, of Theorem 1), up to showing (2.18), which we do at the end of the section.  $\square$

*Proof of (2.18).* We claim that for each  $v \in S^1$

$$(2.19) \quad \mu_{2Q}(\{x \in \mathbb{R}^2 : |T_v(g)(x)| > \gamma\lambda, M_v g(x) < c_0\lambda\}) \leq C e^{-c_1\gamma} \frac{(Cp)^p \|g\|_{L^p(Q)}^p}{\lambda^p}.$$

To prove the claim, assume without loss of generality  $v = (1, 0)$ . Denote by  $g_{x_2}(\cdot)$  the function  $x_1 \mapsto g(x_1, x_2)$ . Then

$$\begin{aligned} & \mu_{2Q}(\{x \in \mathbb{R}^2 : |T_v(g)(x)| > \gamma\lambda, M_v g(x) < c_0\lambda\}) \\ & \leq \frac{1}{4|Q|} \int_{\mathbb{R}} |\{x_1 \in \mathbb{R} : |T(g_{x_2})(x_1)| > \gamma\lambda, M_v g_{x_2}(x_1) < c_0\lambda\}| dx_2 \\ & \leq \frac{1}{4|Q|} \int_{\mathbb{R}} C e^{-c_1\gamma} |\{x_1 \in \mathbb{R} : |T_{\star}(g_{x_2})(x_1)| > \lambda\}| dx_2 \\ & \leq \frac{C}{|Q|} \int_{\mathbb{R}} (Cp)^p e^{-c_1\gamma} \frac{\|g_{x_2}(\cdot)\|_p^p}{\lambda^p} dx_2 \leq \frac{C}{|Q|} e^{-c_1\gamma} \frac{(Cp)^p \|g\|_p^p}{\lambda^p} = C e^{-c_1\gamma} \frac{(Cp)^p \|g\|_{L^p(Q)}^p}{\lambda^p}. \end{aligned}$$

where the third to last inequality is an application of Proposition 2.2, and last line follows from the bound of Proposition 2.1 for  $T_{\star}$  on  $L^p(x_1)$ . Now, using (2.19) for each  $T_v$ , we have

$$\begin{aligned} & \mu_{2Q}(\{x \in \mathbb{R}^2 : T_{\mathbf{V}}g(x) > (c_1)^{-1}\lambda \log N, M_{\mathbf{V}}f(x) < c_0\lambda\}) \\ & \leq \sum_{v \in \mathbf{V}} \mu_{2Q}(\{x \in \mathbb{R}^2 : |T_v(g)(x)| > (c_1)^{-1}\lambda \log N, M_v g(x) < c_0\lambda\}) \\ & \leq C N e^{-\log N} \frac{(Cp)^p \|g\|_{L^p(Q)}^p}{\lambda^p} \leq \frac{(Cp)^p}{\lambda^p} \|g\|_{L^p(Q)}^p, \end{aligned}$$

which completes the proof of (2.18).  $\square$

### 3. LACUNARY $\mathbf{V}$ : PROOF OF THEOREM 2

We need to prove that for a set  $\mathbf{V}$  of  $N$  directions which is lacunary (see Definition 1.2)

$$\|T_{\mathbf{V}}f\|_p \leq C_p \sqrt{\log N} \|f\|_p, \quad 1 < p < \infty$$

In the related paper [6], the  $L^2$ -estimate of the maximal multiplier (1.2) is reduced (with a  $\sqrt{\log N}$  loss) to the  $L^2$  bound for the corresponding square function, via an application of the Chang-Wilson-Wolff inequality [3]. See also [12], where the same method had been first applied to the study of maximal one-dimensional multipliers, and [7]. We apply the same reduction, and then show that the resulting square function is bounded on all  $L^p$ ,  $1 < p < \infty$  uniformly in the cardinality of  $\mathbf{V}$  (see (3.3)).

**3.1. Reduction to a square function estimate.** Let  $\Phi$  be a Schwartz function supported on  $[1/2, 2]$  and such that

$$\sum_{k \in \mathbb{Z}} \Phi(2^{-k}\xi) = 1, \quad \forall \xi \in \mathbb{R} \setminus \{0\}.$$

For a function  $f \in C_0^\infty(\mathbb{R}^2)$ , set  $\widehat{S_k f}(\xi) = \hat{f}(\xi)\Phi(2^{-k}|\xi|)$ , and note that

$$(3.1) \quad \|f\|_p \sim \left\| \left( \sum_{k \in \mathbb{Z}} |S_k f|^2 \right)^{\frac{1}{2}} \right\|_p, \quad 1 < p < \infty.$$

We summarize the reduction to a square function in the following proposition; the proof is contained in the reference [12], see also [6].

**Proposition 3.1.** *Let  $P_i$ ,  $i = 1, \dots, N$  be linear (uniformly) bounded multiplier operators on  $\mathbb{R}^2$ . Denote by  $\mathcal{M}_\beta = M_\beta \circ M_\beta \circ M_\beta$ , where  $M_\beta f := (M[f^\beta])^{\frac{1}{\beta}}$ . Then, for each  $1 < p < \infty$  and each  $\varepsilon > 0$  we have*

$$(3.2) \quad \left\| \sup_{i=1, \dots, N} |P_i f| \right\|_p \lesssim_{\varepsilon, p} (\log N)^{\frac{1}{2}} \left\| \left( \sum_{k \in \mathbb{Z}} \left( \mathcal{M}_{1+\varepsilon} \left[ \sup_{i=1, \dots, N} |P_i(S_k f)| \right] \right)^2 \right)^{\frac{1}{2}} \right\|_p.$$

By means of Proposition 3.1, and a subsequent use of the Fefferman-Stein inequality (with  $1 + \varepsilon < p$ ) we get

$$\left\| \left( \sum_{k \in \mathbb{Z}} \mathcal{M}_{1+\varepsilon}[g_k]^2 \right)^{\frac{1}{2}} \right\|_p \lesssim_p \left\| \left( \sum_{k \in \mathbb{Z}} |g_k|^2 \right)^{\frac{1}{2}} \right\|_p,$$

followed by (3.1), we reduce Theorem 2 to the proof of the estimate

$$(3.3) \quad \left\| \left( \sum_{k \in \mathbb{Z}} |T_{\mathbf{V}}(S_k f)|^2 \right)^{\frac{1}{2}} \right\|_p \lesssim_p \|f\|_p, \quad 1 < p < \infty.$$

Let us recall once again that the easy case  $p = 2$  of (3.3) was proved in [6].

**3.2. Proof of the square function estimate (3.3).** We can harmlessly assume that the multiplier  $m$  is supported on  $[0, \infty)$  and that  $\hat{f}$  is supported in the first quadrant. We take the node of  $\mathbf{V}$  to be  $v_\infty = (1, 0)$  and order  $\mathbf{V} = \{v_j = e^{2\pi i \theta(v_j)}\}$  counterclockwise.

Since  $\theta(v_\infty) = 0$ , (1.9) reads  $|\theta(v_{j+1})| \leq 2^{-1}|\theta(v_j)|$ ,  $j \geq 1$ . By eventually splitting  $\mathbf{V}$  into the three subcollections  $\{v_j \in \mathbf{V} : j \bmod 3 = i\}$ ,  $i \in \{0, 1, 2\}$ , we can work under the stronger separation property

$$(3.4) \quad |\theta(v_{j+1})| \leq 2^{-3}|\theta(v_j)|, \quad j \geq 1.$$

We denote  $\delta = |\theta(v_1)|$ . It is clear that for each interval  $I_\ell := (2^{-3(\ell+1)}\delta, 2^{-3\ell}\delta]$ ,  $\ell \geq 1$ , there is at most one  $j(\ell)$  such that  $\theta(v_{j(\ell)}) \in I_\ell$ . Whenever  $I_\ell \cap \{\theta(v) : v \in \mathbf{V}\} = \emptyset$  for (at least two consecutive)  $\ell, \ell + 1, \dots, \ell + m$ , we add to  $\mathbf{V}$  the  $m$  directions with angles  $\theta \in \{-2^{-3\ell}, \dots, -2^{-3(\ell+m-1)}\}$ . We continue to call  $\mathbf{V}$  the resulting collection and re-index the  $v_j$  so that they are ordered counterclockwise. By doing so, we have obtained that

$$(3.5) \quad 2^{-3(j+1)}\delta \leq \delta_j := |\theta(v_{j+1}) - \theta(v_j)| \leq 2^{-3j}\delta, \quad j \geq 1.$$

It is handy to define  $w_j = \theta(v_j) + \frac{1}{4}$ , since the assumption on the support of the multiplier  $m$  implies that for  $\xi = |\xi|e^{2\pi i\theta(\xi)}$  in the first quadrant

$$(3.6) \quad m(\xi \cdot v_j) \neq 0 \implies \theta(\xi) \in (0, w_j).$$

Accordingly, define the intervals

$$\alpha_0 = (0, w_1], \quad \alpha_j = (w_j, w_{j+1}], \quad j = 1, \dots, \#\mathbf{V} - 1, \quad \alpha_{\#\mathbf{V}} = (w_{\#\mathbf{V}}, w_\infty].$$

For an interval  $I = (\alpha_\ell, \alpha_r) \subset [0, 1]$ , define the frequency cutoff to the cone  $\{e^{2\pi i\theta} : \theta \in I\}$

$$\widehat{G_I f}(\xi) = \hat{f}(\xi) \mathbf{1}_I(\theta(\xi)), \quad \xi = |\xi|e^{2\pi i\theta(\xi)}.$$

We simply write  $G_j$  for  $G_{\alpha_j}$ . The bulk of the proof is contained in the next two results. In the first lemma, which is proven in Subsection 3.3, a pointwise bound on  $T_{\mathbf{V}}(S_k f)$  is given, in terms (essentially) of the maximal function along the direction  $v_\infty$ .

**Lemma 3.2.** *Define*

$$(3.7) \quad f_{\text{ev}} = \sum_j G_{2j} f, \quad f_{\text{odd}} = \sum_j G_{2j+1} f,$$

and adopt the notation (here  $v^\perp$  is the direction orthogonal to  $v$ )

$$M_v^{\text{bi}} f(x) = \sup_{t,s>0} \frac{1}{4ts} \int_{-t}^t \int_{-s}^s |f(x + \tau v + \sigma v^\perp)| \, d\tau d\sigma, \quad x \in \mathbb{R}^2,$$

for the bi-parameter maximal function in the coordinates of  $v$ . Then

$$(3.8) \quad T_{\mathbf{V}}(S_k f)(x) \lesssim M_{v_\infty}^{\text{bi}}[S_k f_{\text{ev}}](x) + M_{v_\infty}^{\text{bi}}[S_k f_{\text{odd}}](x), \quad x \in \mathbb{R}^2, \quad k \in \mathbb{Z}.$$

The second lemma deals with the boundedness of the rough cutoff multiplier  $\sum_j \varepsilon_j G_j$ . The proof is given in Subsection 3.4.

**Lemma 3.3.** *We have the estimate*

$$(3.9) \quad \sup_{\varepsilon_j=0,\pm 1} \left\| \sum_j \varepsilon_j G_j f \right\|_p \lesssim \|f\|_p, \quad 1 < p < \infty.$$

*Conclusion of the proof of (3.3).* Lemma 3.3 immediately implies the bounds

$$(3.10) \quad \|f_{\text{ev}}\|_p, \|f_{\text{odd}}\|_p \lesssim \|f\|_p, \quad 1 < p < \infty.$$

We conclude, using the Fefferman-Stein inequality on the vector valued  $M_{v_\infty}^{\text{bi}}$ , that

$$\begin{aligned} & \left\| \left( \sum_{k \in \mathbb{Z}} |T_{\mathbf{V}}(S_k f)|^2 \right)^{\frac{1}{2}} \right\|_p \\ & \lesssim \left\| \left( \sum_{k \in \mathbb{Z}} [M_{v_\infty}^{\text{bi}}[S_k f_{\text{ev}}]]^2 \right)^{\frac{1}{2}} \right\|_p + \left\| \left( \sum_{k \in \mathbb{Z}} [M_{v_\infty}^{\text{bi}}[S_k f_{\text{odd}}]]^2 \right)^{\frac{1}{2}} \right\|_p \\ & \lesssim \left\| \left( \sum_{k \in \mathbb{Z}} |S_k f_{\text{ev}}|^2 \right)^{\frac{1}{2}} \right\|_p + \left\| \left( \sum_{k \in \mathbb{Z}} |S_k f_{\text{odd}}|^2 \right)^{\frac{1}{2}} \right\|_p \\ & \lesssim \|f_{\text{ev}}\|_p + \|f_{\text{odd}}\|_p \lesssim \|f\|_p. \end{aligned}$$

This completes the proof of (3.3), and in turn, of Theorem 2.  $\square$

**3.3. Proof of Lemma 3.2.** The estimate (3.8) is an easy consequence of

$$(3.11) \quad |T_{v_j}[S_k f_{\text{ev}}](x)| \lesssim M_{v_\infty}^{\text{bi}}[S_k f_{\text{ev}}](x), \quad \forall x \in \mathbb{R}^2, v_j \in \mathbf{V},$$

and of the analogue estimate for  $S_k f_{\text{odd}}$ . Observe that  $T_{v_{2j}}[S_k f_{\text{ev}}] = T_{v_{2j-1}}[S_k f_{\text{ev}}]$ , since  $\widehat{f_{\text{ev}}}$  is zero on the cone bordered by  $w_{2j-1}$  and  $w_{2j}$ . Thus it suffices to consider odd  $j$ . We now show (3.11). Let  $j$  be fixed throughout. We will omit some dependence on  $j$  in the notation. Denote by

$$(\eta_1, \eta_2) := \{tw_j : t > 0\} \cap \{|\xi| = 2^k\}, \quad (\zeta_1, \zeta_2) := \{tw_{j+1} : t > 0\} \cap \{|\xi| = 2^{k+1}\},$$

As a consequence of (3.4), we have that

$$0 < \zeta_1 = 2^{k+1} \cos(w_{j+1}) \leq 2^{k+1} |\theta(v_{j+1})| < 2^{k-2} |\theta(v_j)| \leq 2^k \cos(w_j) = \eta_1 \sim 2^k \delta 8^{-j}$$

Now, let  $\psi_j$  be a positive Schwartz function on  $\mathbb{R}$  which is equal to 1 on  $[2^k 8^{-j-1} \delta, 2^k]$  and vanishes outside of  $[2^k 8^{-j-2} \delta, 2^{k+1}]$ . The idea is that we can insert the vertical frequency cutoff given by  $\psi_j(\xi_1)$  harmlessly, because (see (3.6))

$$\widehat{[S_k f_{\text{ev}}]}(\xi) m(\xi \cdot v_j) \neq 0 \implies \xi_1 \geq \eta_1 \implies \psi_j(\xi_1) = 1,$$

and therefore  $\widehat{T_{v_j}[S_k f_{\text{ev}}]}(\xi) = \widehat{S_k f_{\text{ev}}}(\xi) \mu_j(\xi)$ , where we set  $\mu_j(\xi) = \Phi(2^{-k} |\xi|) \psi_j(\xi_1) m(\xi \cdot v_j)$ . The multiplier  $\mu_j$  is easily seen to satisfy

$$|\partial_{\xi_1}^\alpha \partial_{\xi_2}^\beta \mu_j(\xi_1, \xi_2)| \lesssim 8^{j\alpha} \delta^{-\alpha} 2^{-k\alpha} 2^{-k\beta}, \quad 0 \leq \alpha, \beta \leq \kappa,$$

for some constant  $\kappa > 3$ . Thus  $K_j = \check{\mu}_j$  obeys the bound

$$|K_j(x_1, x_2)| \lesssim 2^{2k} \delta 8^{-j} |1 + 2^k \delta 8^{-j} |x_1| + 2^k |x_2|||^{-3}$$

This implies

$$|T_{v_j}[S_k f_{\text{ev}}]|(x_1, x_2) \lesssim |S_k f_{\text{ev}} * K_j|(x_1, x_2) \lesssim M_{v_\infty}^{\text{bi}}[S_k f_{\text{ev}}](x_1, x_2),$$

as claimed in (3.11). The proof of Lemma 3.2 is complete.

**3.4. Proof of Lemma 3.3.** We need a smooth analogue of  $G_I$ . Let  $\beta$  be a positive Schwartz function with support in  $[-\frac{1}{2}, \frac{3}{2}]$  and such that  $\sum_{\ell \in \mathbb{Z}} \beta(t - \ell) = \mathbf{1}_{\mathbb{R}}(t)$ ; define the multiplier operator

$$\widehat{G_I^s f}(\xi) := \hat{f}(\xi) \beta\left(\frac{\theta(\xi) - \alpha_\ell}{|I|}\right), \quad \xi = |\xi| e^{2\pi i \theta(\xi)} \in \mathbb{R}^2.$$

Let us cover each  $\alpha_j$  by 128 intervals of equal length  $\alpha_{j,\ell} = (w_j + (\ell - 1) \frac{\delta_j}{128}, w_j + \ell \frac{\delta_j}{128}]$ , and denote  $G_{\alpha_{j,\ell}}^s$  simply by  $G_{j,\ell}^s$ . By the definition of  $\beta$ , we have the identity

$$(3.12) \quad G_j f = G_j(G_j^s f), \quad G_j^s f := \sum_{\ell=1}^{128} G_{j,\ell}^s f.$$

The frequency support of  $G_{j,\ell}^s$  is contained in  $2\alpha_{j,\ell}$  (the interval with same center as  $\alpha_{j,\ell}$  and twice the length). By (3.5),  $\delta_j$  and  $\delta_{j+1}$  are within a factor of 16, thus

$$\sum_j \sum_{\ell=1}^{128} \mathbf{1}_{2\alpha_{j,\ell}}(t) \leq 32.$$

The above observations and the smoothness of  $\beta$  imply that  $G = \sum_{j,\ell} \varepsilon_{j,\ell} G_{j,\ell}^s$  is a (two-dimensional) Hörmander-Mihlin multiplier, and therefore

$$(3.13) \quad \sup_{\varepsilon_j=0,\pm 1} \left\| \sum_{j,\ell} \varepsilon_{j,\ell} G_{j,\ell}^s h \right\|_p \lesssim_p \|h\|_p, \quad 1 < p < \infty;$$

see [28, Theorem IV.3]. By randomization, we get the square function inequality

$$(3.14) \quad \left\| \left( \sum_{j,\ell} |G_{j,\ell}^s(h)|^2 \right)^{\frac{1}{2}} \right\|_p \lesssim \|h\|_p, \quad 1 < p < \infty.$$

We will need the vector-valued inequality of the Proposition below, which is a variant of the results of [5] and its refinement [11, Theorem VI.8.1]. The full range  $1 < p < \infty$  is obtained by combining the method used in [11] with the sharp weighted bound for the Hilbert transform in terms of the  $A_2$  characteristic of the weight [26].

**Proposition 3.4.** *For each  $1 < p < \infty$ , we have that*

$$(3.15) \quad \left\| \left( \sum_j |G_j[h_j]|^2 \right)^{\frac{1}{2}} \right\|_p \lesssim \left\| \left( \sum_j |h_j|^2 \right)^{\frac{1}{2}} \right\|_p.$$

Recall the identity (3.12). Applying (3.15) for  $h_j = G_j^s g$ , and subsequently using (3.14), we get that for each  $g \in L^2 \cap L^p$

$$\begin{aligned} \left\| \left( \sum_j |G_j g|^2 \right)^{\frac{1}{2}} \right\|_p &= \left\| \left( \sum_j |G_j[G_j^s g]|^2 \right)^{\frac{1}{2}} \right\|_p \lesssim \left\| \left( \sum_j |G_j^s g|^2 \right)^{\frac{1}{2}} \right\|_p \\ &\lesssim \left\| \left( \sum_{j,k} |G_{j,k}^s g|^2 \right)^{\frac{1}{2}} \right\|_p \lesssim \|g\|_p \end{aligned}$$

When  $p = 2$ , the first and the last quantities in the above display coincide, so that polarization actually gives the equivalence

$$(3.16) \quad c_p \|g\|_p \leq \left\| \left( \sum_j |G_j g|^2 \right)^{\frac{1}{2}} \right\|_p \leq C_p \|g\|_p, \quad 1 < p < \infty.$$

Applying the left inequality in (3.16) to  $g = \sum_j \varepsilon_j G_j f$  then yields

$$\|g\|_p \lesssim \left\| \left( \sum_j |G_j g|^2 \right)^{\frac{1}{2}} \right\|_p \leq \left\| \left( \sum_j |G_j f|^2 \right)^{\frac{1}{2}} \right\|_p \lesssim \|f\|_p, \quad 1 < p < \infty.$$

where the last step follows from the right inequality in (3.16). The last display is exactly the assertion we had to prove.

#### 4. GRIDS, ADAPTEDNESS, AND A PRODUCT JOHN-NIRENBERG INEQUALITY

In this section, we define product size and establish a version of the (product) John-Nirenberg inequality, Proposition 4.1. Most of these results are well known from previous literature. Let us start with some definitions.

4.1. **Grids.** A one-dimensional *grid*  $\mathcal{G}$  is a collection of intervals with the property that

$$(4.1) \quad I, I' \in \mathcal{G} \implies I \cap I' \in \{\emptyset, I, I'\}.$$

For  $i \in \{0, 1, 2\}$ ,  $\mathcal{D}^i = \{2^j(\ell + \frac{i(-1)^j}{3} + [0, 1) : j, \ell \in \mathbb{Z})\}$  are three (dyadic) grids with the property that for each interval  $J \subset \mathbb{R}$  there exists some  $I \in \mathcal{D} \cup \mathcal{D}^1 \cup \mathcal{D}^2$  with  $J \subset I \subset 9J$ . For simplicity, we write  $\mathcal{D}$  for  $\mathcal{D}^0$ ; for a scale  $\text{scl} \in 2^{\mathbb{Z}}$ ,  $\mathcal{D}_{\text{scl}}^i = \{I \in \mathcal{D}^i : |I| = \text{scl}\}$ .

A grid  $\mathcal{G}$  is said to be  $K$ -sparse if for every two  $I, J \in \mathcal{G}$

$$(4.2) \quad |I| < |I'| \implies |I| \leq 2^K |I'|, \quad |I| = |I'| \implies \text{dist}(I, I') \geq 2^K |I'|.$$

Right from the definition, if  $\mathcal{G}$  is  $(K+10)$ -sparse and  $I, I' \in \mathcal{G}$ , with  $|I| = |I'|$ , the dilates  $2^{K+4}I$  and  $2^{K+4}I'$  do not intersect. On the other hand,  $I, I' \in \mathcal{G}, I \subset I' \implies 2^{K+4}I \subset 2I'$ . Using this property, it follows that if  $\mathcal{G}$  is  $(K+10)$ -sparse,

$$(4.3) \quad \forall I \in \mathcal{G} \exists J(I) \in \mathcal{D}^0 \cup \mathcal{D}^1 \cup \mathcal{D}^2 \text{ with } 2^K I \subset J(I) \subset 9 \cdot 2^K I,$$

and  $|I| = |I'| \implies J(I) \cap J(I') = \emptyset, I \subset I' \implies J(I) \subset J(I')$ . Now, observe that every dyadic grid  $\mathcal{G}$  can be written as the disjoint union of  $K2^K$   $K$ -sparse dyadic grids. It follows that  $2^K \mathcal{G} := \{2^K I : I \in \mathcal{G}\}$  can be written as the union of  $3 \cdot K2^K$  subsets, each of which fits into a dyadic grid as described in (4.3).

4.2.  **$L^2$ -adaptedness.** Let  $R = I_R \times J_R$  be a rectangle in  $\mathbb{R}^2$  with orientation  $(e_R, e_R^\perp)$ . We denote by  $c(R)$  its center and by  $R^{(k)}$  the rectangle  $2^k R$ . Let  $K$  be a large positive constant. We say that a Schwartz function  $\psi_R$  is  $L^2$ -adapted to the rectangle  $R$  if for all  $\alpha, \beta > 0$  there exists a constant  $C_{\alpha, \beta}$  such that, in the coordinates  $(e_R, e_R^\perp)$ ,

$$(4.4) \quad |\partial_{x_1}^\alpha \partial_{x_2}^\beta \psi_R(x_1, x_2)| \leq \frac{C_{\alpha, \beta}}{|I_R|^{\alpha+\frac{1}{2}} |J_R|^{\beta+\frac{1}{2}}} \left(1 + \frac{|x_1 - c(R)_1|}{|I_R|} + \frac{|x_2 - c(R)_2|}{|J_R|}\right)^{-K}.$$

Let  $\mathcal{R}$  be a collection of rectangles (with possibly different orientations). We say that  $\{\psi_R\}_{R \in \mathcal{R}}$  are  $L^2$ -adapted to  $\mathcal{R}$  if each  $\psi_R$  is  $L^2$ -adapted with same the constants  $C_{\alpha, \beta}, K$ .

If  $\{\psi_R\}_{R \in \mathcal{R}}$  are  $L^2$ -adapted and for each  $R$  we have  $\int_{\mathbb{R}^2} \psi_R = 0$ , we say that  $\{\psi_R\}_{R \in \mathcal{R}}$  are *wave packets* adapted to  $\mathcal{R}$ . If the stronger condition

$$(4.5) \quad \int_{\{x+te_R:t \in \mathbb{R}\}} \psi_R = \int_{\{x+te_R^\perp:t \in \mathbb{R}\}} \psi_R = 0, \quad \forall x \in \mathbb{R}^2$$

holds, we say that  $\{\psi_R\}_{R \in \mathcal{R}}$  are product wave packets adapted to  $\mathcal{R}$ .

4.3. **A product John-Nirenberg inequality.** Let  $\mathcal{R}$  be a collection of rectangles  $R = I_R \times J_R$  with fixed orientation. Assume that  $\mathcal{I}_{\mathcal{R}} := \{I_R : R \in \mathcal{R}\}$  and  $\mathcal{J}_{\mathcal{R}} := \{J_R : R \in \mathcal{R}\}$  are each (respectively) subsets of a dyadic grid  $\mathcal{I}$  (resp.  $\mathcal{J}$ ). We will use the notation

$$\text{sh}(\mathcal{R}) = \bigcup_{R \in \mathcal{R}} R$$

---

<sup>1</sup>By  $CJ$  we mean the interval with same center as  $J$  and  $C$  times longer. If  $R = I \times J$  is a rectangle,  $CR = CI \times CJ$ .

for the *shadow* of such a collection. Given a collection of complex coefficients  $B = \{b_R\}_{R \in \mathcal{R}}$  and functions  $\{\psi_R\}_{R \in \mathcal{R}}$ , define

$$(4.6) \quad \mathbf{B}_{\mathcal{R}'}(x) = \sum_{R \in \mathcal{R}'} b_R \psi_R(x), \quad \mathcal{R}' \subset \mathcal{R}.$$

Let us introduce

$$(4.7) \quad \text{size}(B) = \sup_{\mathcal{R}' \subset \mathcal{R}} \left( \frac{1}{|\text{sh}(\mathcal{R}')|} \sum_{R \in \mathcal{R}'} |b_R|^2 \right)^{\frac{1}{2}}$$

In the following proposition, we show how size is akin to the product BMO norm.

**Proposition 4.1.** *Let  $\{\psi_R\}_{R \in \mathcal{R}}$  be product wave packets adapted to  $\mathcal{R}$ , and with*

$$(4.8) \quad \text{supp } \psi_R \subset R.$$

*We have the estimates*

$$(4.9) \quad \|\mathbf{B}_{\mathcal{R}}\|_p \leq Cp^2 \text{size}(B) |\text{sh}(\mathcal{R})|^{\frac{1}{p}}, \quad 2 \leq p < \infty,$$

$$(4.10) \quad \left| \{x \in \text{sh}(\mathcal{R}) : |B_{\mathcal{R}}(x)| > \lambda\} \right| \leq C \exp \left( -c \sqrt{\frac{\lambda}{\text{size}(B)}} \right) |\text{sh}(\mathcal{R})|.$$

Subsections 4.4 and 4.5 are devoted to the proofs (respectively) of Proposition 4.1 and of the related product Chang-Wilson Wolff inequality (Proposition 4.5). We can actually show that estimate (4.9) continues to hold even when the restriction (4.8) on the support of the  $\psi_R$  is removed. This will be done in Subsection 4.6.

**Remark 4.2.** We also detail a one parameter version of Proposition 4.1, whose proof follows the same lines. Let  $\mathcal{R}$  be a collection of rectangles as above, with fixed eccentricity  $\frac{|J_R|}{|I_R|} \equiv \text{ecc}$ . Let  $\{\psi_R\}_{R \in \mathcal{R}}$  be wave packets adapted to  $\mathcal{R}$ , and with  $\text{supp } \psi_R \subset R$ . Then

$$(4.11) \quad \left| \{x \in \text{sh}(\mathcal{R}) : |B_{\mathcal{R}}(x)| > \lambda\} \right| \leq C \exp \left( -c \frac{\lambda}{\text{size}(B)} \right) |\text{sh}(\mathcal{R})|.$$

**4.4. Proof of Proposition 4.1.** For simplicity, say  $\mathcal{I} = \mathcal{J} = \mathcal{D}$ . Proposition 4.1 relies upon two results. First, we establish an inequality similar to (4.9) for the related square function

$$(4.12) \quad \mathbf{SB}_{\mathcal{R}'}(x) = \left( \sum_{R \in \mathcal{R}'} \frac{|b_R|^2}{|R|} \mathbf{1}_R(x) \right)^{\frac{1}{2}}, \quad \mathcal{R}' \subset \mathcal{R}.$$

**Lemma 4.3.** *We have the estimate*

$$(4.13) \quad \|\mathbf{SB}_{\mathcal{R}}\|_p \leq Cp \text{size}(B) |\text{sh}(\mathcal{R})|^{\frac{1}{p}}, \quad 2 < p < \infty.$$

Second, we obtain a sharp (with respect to  $p$ ) norm comparison between  $\mathbf{B}_{\mathcal{R}}$  and  $\mathbf{SB}_{\mathcal{R}}$ .

**Lemma 4.4.** *We have the estimate*

$$(4.14) \quad \|\mathbf{B}_{\mathcal{R}}\|_p \leq Cp \|\mathbf{SB}_{\mathcal{R}}\|_p, \quad 2 \leq p < \infty.$$



Thus, (4.9) is obtained by applying Lemmata 4.3 and 4.4. Then, (4.10) follows immediately from (4.9) by extrapolation.

The proofs of Lemma 4.3 and Lemma 4.4 are given at the end of this section. Lemma 4.4 will be obtained as a consequence of the product Chang-Wilson-Wolff inequality of [27], which we detail in Proposition 4.5 in the form of sharp (with respect to  $p$ , as  $p \rightarrow \infty$ )  $L^p$  bounds between  $f$  and its product martingale square function, following the approach of [10]. We recall some notation. For  $I \in \mathcal{D}$ ,  $h_I$  denotes the  $L^2$  normalized Haar wavelet on  $I$ . For a rectangle  $Q = I \times J$ ,  $h_Q$  stands for the tensor product wavelet  $h_I \otimes h_J$ . We introduce the one and two-parameter martingale square functions

$$(4.15) \quad (\Delta_1 f)(x_1) = \left( \sum_{I \in \mathcal{D}} |\langle f, h_I \rangle|^2 \frac{\mathbf{1}_I(x_1)}{|I|} \right)^{\frac{1}{2}}, \quad (\Delta_{12} f)(x) = \left( \sum_{Q \in \mathcal{D} \times \mathcal{D}} |\langle f, h_Q \rangle|^2 \frac{\mathbf{1}_Q(x)}{|Q|} \right)^{\frac{1}{2}}.$$

**Proposition 4.5.** *We have the sharp inequality*

$$(4.16) \quad \|f\|_p \leq Cp \|\Delta_{12} f\|_p, \quad 2 \leq p < \infty.$$

The proof of Proposition 4.5 is outlined in Subsection 4.5.

*Proof of Lemma 4.3.* Let  $p > 2$  and  $q = \frac{p}{2}$ . It then suffices to show that

$$(4.17) \quad \|(\mathbf{SB}_{\mathcal{R}})^2\|_{L^q(\text{sh}(\mathcal{R}))} \leq Cq \|\mathbf{M}_{\mathcal{R}}\|_{q' \rightarrow q', \infty} \text{size}(B)^2 |\text{sh}(\mathcal{R})|^{\frac{1}{q}}, \quad 1 \leq q < \infty.$$

Indeed, the bound  $\|\mathbf{M}_{\mathcal{R}}\|_{q' \rightarrow q', \infty} \leq Cq$  follows immediately via weak-type interpolation of the trivial  $L^\infty$  bound with the endpoint for the strong maximal function on  $\mathbb{R}^2$  [JMZ]

$$(4.18) \quad |\{x \in \text{sh}(\mathcal{R}) : \mathbf{M}_{\mathcal{R}} f(x) > \lambda\}| \leq C \int_{\text{sh}(\mathcal{R})} \frac{|f(x)|}{\lambda} \log \left( e + \frac{|f(x)|}{\lambda} \right) dx.$$

For simplicity, denote  $K_q := \|\mathbf{M}_{\mathcal{R}}\|_{q' \rightarrow q', \infty}$ . To prove (4.17), we assume the following claim.

CLAIM. For each  $\mathcal{R}' \subset \mathcal{R}$  there exists  $\mathcal{R}'' \subset \mathcal{R}'$  with

$$(4.19) \quad |\text{sh}(\mathcal{R}'')| \leq \frac{|\text{sh}(\mathcal{R}')|}{2}, \quad \|(\mathbf{SB}_{\mathcal{R}'}^2)\|_q \leq 2^{\frac{1}{q'}} K_q \text{size}(B)^2 |\text{sh}(\mathcal{R}')|^{\frac{1}{q}} + \|(\mathbf{SB}_{\mathcal{R}''}^2)\|_q.$$

With the claim in hand, we can easily complete the proof. Apply the claim to  $\mathcal{R}' = \mathcal{R}$ ; this yields an  $\mathcal{R}'' := \mathcal{R}_1$  with  $\text{sh}(\mathcal{R}_1) \leq 2^{-1} \text{sh}(\mathcal{R})$ . Apply the claim with  $\mathcal{R}' = \mathcal{R}_1$ . This yields an  $\mathcal{R}'' = \mathcal{R}_2$  with  $\text{sh}(\mathcal{R}_2) \leq 2^{-2} \text{sh}(\mathcal{R})$ . Iterating, one finally gets

$$(4.20) \quad \|(\mathbf{SB}_{\mathcal{R}'}^2)\|_q \leq CK_q \text{size}(B)^2 \sum_{k=0}^{\infty} 2^{-\frac{k}{q}} |\text{sh}(\mathcal{R})|^{\frac{1}{q}} \leq \frac{CK_q}{1 - 2^{-\frac{1}{q}}} \text{size}(B)^2 |\text{sh}(\mathcal{R})|^{\frac{1}{q}}$$

and the proposition follows from the fact that  $1 - 2^{-\frac{1}{q}} \geq cq^{-1}$ .

The claim is proven using duality. By scaling, we can assume  $\text{size}(B) = 1$ , which simplifies the notation. We choose  $g \in L^{q'}$ ,  $\|g\|_{q'} = 1$ , so that  $\|(\mathbf{SB}_{\mathcal{R}'}^2)\|_q = \langle (\mathbf{SB}_{\mathcal{R}'}^2), g \rangle$ , and set

$$\mathcal{R}'' = \{R \in \mathcal{R}' : \mathbb{E}_{\mathcal{R}} g \geq 2^{\frac{1}{q'}} K_q |\text{sh}(\mathcal{R}')|^{-\frac{1}{q'}}\}.$$

By the bound on  $\mathbf{M}_{\mathcal{R}}$ , we have that

$$|\text{sh}(\mathcal{R}'')| \leq \frac{1}{2} |\text{sh}(\mathcal{R}')| \text{ and } \forall R \in \mathcal{R}' \setminus \mathcal{R}'', \mathbb{E}_{\mathcal{R}} g < 2^{\frac{1}{q'}} K_q |\text{sh}(\mathcal{R}')|^{-\frac{1}{q'}}.$$

Therefore

$$\begin{aligned}
\|(\mathbf{SB}_{\mathcal{R}'} )^2\|_q &= \langle (\mathbf{SB}_{\mathcal{R}'} )^2, g \rangle \\
&= \left( \sum_{R \in \mathcal{R}' \setminus \mathcal{R}''} |b_R|^2 \mathbb{E}_{\mathcal{R}} g \right) + \|(\mathbf{SB}_{\mathcal{R}''})^2\|_q \|g\|_{q'} \\
&\leq \left( \sup_{R \in \mathcal{R}' \setminus \mathcal{R}''} \mathbb{E}_{\mathcal{R}} g \right) \left( \sum_{R \in \mathcal{R}' \setminus \mathcal{R}''} |b_R|^2 \right) + \|(\mathbf{SB}_{\mathcal{R}''})^2\|_q \\
&\leq CK_q |\text{sh}(\mathcal{R}')|^{1-\frac{1}{q'}} + \|(\mathbf{SB}_{\mathcal{R}'} )^2\|_q,
\end{aligned}$$

and this completes the proof of the claim, and in turn, of Lemma 4.3.  $\square$

*Proof of Lemma 4.4.* The lemma is an immediate consequence of Proposition 4.5 applied to  $f = \mathbf{B}_{\mathcal{R}}$ , and of the pointwise bound

$$(4.21) \quad (\Delta_{12} \mathbf{B}_{\mathcal{R}})(x) \leq C \mathbf{SB}_{\mathcal{R}}(x), \quad x \in \text{sh}(\mathcal{R}).$$

Let us prove (4.21), by computing  $\Delta_{12} \mathbf{B}_{\mathcal{R}}$  explicitly. Observe that

$$(4.22) \quad Q \not\subset R \implies \langle \psi_R, h_Q \rangle = 0.$$

Indeed, if  $Q = I_Q \times J_Q$  is not contained in  $R = I_R \times J_R$ , and  $Q \cap R \neq \emptyset$ , we have that either  $I_R \subsetneq I_Q$ , or  $J_R \subsetneq J_Q$  (or both). Then (4.22) is a consequence of  $\psi_R$  having mean zero (see (4.5)), and  $h_Q$  being constant, along each horizontal (resp. vertical) line. Moreover, again as a consequence of (4.5) and of the smoothness of  $\psi_R$ ,

$$(4.23) \quad Q \subset R \implies |\langle \psi_R, h_Q \rangle| \lesssim \left( \frac{|Q|}{|R|} \right)^{\frac{3}{2}}.$$

Therefore

$$\begin{aligned}
[(\Delta_{12} \mathbf{B}_{\mathcal{R}})(x)]^2 &= \sum_{Q \in \mathcal{D} \times \mathcal{D}} \left( \sum_{R \in \mathcal{R}: R \supseteq Q} b_R \langle \psi_R, h_Q \rangle \right)^2 \frac{\mathbf{1}_Q(x)}{|Q|} \\
&\lesssim \sum_{Q \in \mathcal{D} \times \mathcal{D}} \left( \sum_{R \in \mathcal{R}: R \supseteq Q} \frac{b_R}{\sqrt{|R|}} \frac{|Q|}{|R|} \right)^2 \mathbf{1}_Q(x) \\
&\lesssim \sum_{Q \in \mathcal{D} \times \mathcal{D}} \left( \sum_{R \in \mathcal{R}: R \supseteq Q} \frac{|b_R|^2 |Q|}{|R|} \right) \left( \sum_{R \in \mathcal{R}: R \supseteq Q} \frac{|Q|}{|R|} \right) \mathbf{1}_Q(x) \\
&\lesssim \sum_{R \in \mathcal{R}} \frac{|b_R|^2}{|R|} \left( \sum_{k, \ell \geq 0} \sum_{\substack{Q \subseteq R \\ |I_Q|=2^{-k}|I_R| \\ |J_Q|=2^{-\ell}|J_R|}} 2^{-k} 2^{-\ell} \mathbf{1}_Q(x) \right) \lesssim \sum_{R \in \mathcal{R}} \frac{|b_R|^2}{|R|} \mathbf{1}_R(x) = [\mathbf{SB}_{\mathcal{R}}(x)]^2,
\end{aligned}$$

which is exactly (4.21).  $\square$

**4.5. Proof of Proposition 4.5.** We follow the approach of [10]. The bound (4.16) is obtained as a consequence of the vector-valued inequality for the one parameter square function

$$(4.24) \quad \left\| \| (g_I) \|_{\ell^2(I)} \right\|_{L_{x_2}^p} \leq Cp^{\frac{1}{2}} \left\| \Delta_2 g_I \right\|_{\ell^2(I)} \Big\|_{L_{x_2}^p}, \quad 2 \leq p < \infty.$$

Consider the  $\ell^2$ -valued function  $F_I(x_1, x_2) = \langle f(\cdot, x_2), h_I \rangle h_I(x_1)$ . Then  $\Delta_{12}f = \|\Delta_2 F_I\|_{\ell^2(I)}$ , so that

$$\begin{aligned} \|\Delta_{12}f\|_p &= \left\| \left\| (\Delta_2 F_I) \right\|_{\ell^2(I)} \right\|_p = \left\| \left\| \left\| \Delta_2 F_I(x_1, \cdot) \right\|_{\ell^2(I)} \right\|_{L_{x_2}^p} \right\|_{L_{x_1}^p} \\ &\geq C^{-1} p^{-\frac{1}{2}} \left\| \left\| F_I(x_1, \cdot) \right\|_{\ell^2(I)} \right\|_{L_{x_2}^p} \Big|_{L_{x_1}^p} = C^{-1} p^{-\frac{1}{2}} \left\| \left\| F_I(x_1, \cdot) \right\|_{\ell^2(I)} \right\|_{L_{x_1}^p} \Big|_{L_{x_2}^p} \\ &= C^{-1} p^{-\frac{1}{2}} \left\| \left\| \Delta_1 f(\cdot, x_2) \right\|_{L_{x_1}^p} \right\|_{L_{x_2}^p} \geq C^{-2} p^{-1} \left\| \left\| f(\cdot, x_2) \right\|_{L_{x_1}^p} \right\|_{L_{x_2}^p} = C^{-2} p^{-1} \|f\|_p. \end{aligned}$$

We used the vector-valued bound (4.24) to get to the second line, and the scalar valued bound to obtain the last inequality. We have thus proved Proposition 4.5, up to showing that (4.24) holds. This is done by applying the Rubio De Francia trick to the weighted inequality

$$(4.25) \quad \int_{\mathbb{R}} \|f_k\|_{\ell^2(k)}^2 w(x_2) dx_2 \leq C[w]_{A_1} \int_{\mathbb{R}} \|\Delta_2 f_k\|_{\ell^2(k)}^2 w(x_2) dx_2,$$

which is obviously equivalent to the scalar case. In turn, the scalar case of (4.25) follows from the Chang-Wilson-Wolff good- $\gamma$  inequality [3]

$$\left| \left\{ x_2 : \sup_{x_2 \in I} |\mathbb{E}_I f| > 2\lambda, \Delta_2 f(x_2) \leq \gamma\lambda \right\} \right| \leq C e^{-\frac{c}{\gamma^2}} \left| \left\{ x_2 : \sup_{x_2 \in I} |\mathbb{E}_I f| > \lambda \right\} \right|.$$

One chooses  $\frac{1}{\gamma^2} = c[w]_{A_1}$  and recalls the following property of  $A_1$  weights: there exists  $c > 0$  such that

$$E \subset I, |E| \leq e^{-c[w]_{A_1}} |I| \implies w(E) \leq \frac{1}{8} w(Q).$$

This yields a weighted good-lambda inequality whose integration gives (the scalar case of) (4.25). See [10] for details.

**4.6. Dropping the compact support assumption.** Let  $\mathcal{R}$  be as above, and now let  $\{\psi_R\}_{R \in \mathcal{R}}$  be product wave packets  $L^2$ -adapted to  $\mathcal{R}$  for which we no longer assume (4.8). Let us prove that (4.9) continues to hold. The main tool is the following lemma, which is a variant of [24, Lemma 3.1].

**Lemma 4.6.** *Let  $\psi_R$  be a function  $L^2$  adapted to the rectangle  $R$  and having mean zero along each line parallel to the coordinate axes of  $R$ , as in (4.5). Then*

$$(4.26) \quad \psi_R(x_1, x_2) = \sum_{k \geq 0} 2^{-100k} \psi_{R^{(k)}}(x_1, x_2),$$

where for each  $k \geq 0$ ,  $\{\psi_{R^{(k)}}\}$  is a collection of product wave packets adapted to  $\{R^{(k)} : R \in \mathcal{R}\}$ , with  $\text{supp } \psi_{R^{(k)}} \subset R^{(k)}$ . The adaptation constants depend only on the adaptation constants of  $\psi_R$  (in particular, they do not depend on  $k$ ).

We apply Lemma 4.6 to each  $\psi_R$ , so that

$$(4.27) \quad B_{\mathcal{R}}(x) = \sum_{R \in \mathcal{R}} b_R \psi_R(x) = \sum_{k \geq 0} 2^{-100k} \sum_{\rho \in \mathcal{R}^{(k)}} b_{\rho} \psi_{\rho}(x) = \sum_{k \geq 0} 2^{-100k} B_{\mathcal{R}^{(k)}}^{(k)}(x).$$

where  $\mathcal{R}^{(k)} := \{\rho = R^{(k)} : R \in \mathcal{R}\}$  and we defined  $B^{(k)} = \{b_{\rho} := b_R : \rho = R^{(k)} \in \mathcal{R}^{(k)}\}$ . Then, it is easy to see that  $\text{size}(B^{(k)}) \leq \text{size}(B)$ . As discussed in Subsection 4.1,  $\mathcal{R}^{(k)}$  can

be split into  $Ck^2 2^{2k}$  subcollections of rectangles whose horizontal and vertical sides fit into a single dyadic grid. Thus, applying (4.9) for each of these subsets,

$$\|B_{\mathcal{R}^{(k)}}^{(k)}\|_p \leq Ck^2 2^{2k} p^2 \text{size}(B^{(k)}) |\text{sh}(\mathcal{R}^{(k)})|^{\frac{1}{p}}.$$

Finally, using (4.27), we conclude that

$$(4.28) \quad \begin{aligned} \|B_{\mathcal{R}}\|_p &\leq \sum_{k \geq 0} 2^{-100k} \|B_{\mathcal{R}^{(k)}}^{(k)}\|_p \leq Cp^2 \left( \sum_{k \geq 0} 2^{-90k} \text{size}(B^{(k)}) |\text{sh}(\mathcal{R}^{(k)})|^{\frac{1}{p}} \right) \\ &\leq Cp^2 \text{size}(B) |\text{sh}(\mathcal{R})|^{\frac{1}{p}}, \end{aligned}$$

as claimed.

## 5. THE TIME-FREQUENCY PHASE PLANE

We briefly recall (see [23]) how to produce a discretization of the multiplier operator

$$\widehat{G_\alpha f}(|\xi| e^{2\pi i \theta(\xi)}) = \hat{f}(\xi) \mathbf{1}_\alpha(\theta(\xi)).$$

where  $\alpha = (0, \alpha)$  is a subinterval of  $(0, 1)$ .

Following [23], we start by decomposing the frequency plane into a union of annular sectors. For  $\mathbf{ann} \in 2^{\mathbb{Z}}$  and  $\omega \in \mathcal{D}$ ,  $\omega \in [0, 1]$  define

$$\Omega_{\mathbf{ann}, \omega} = \{\xi = |\xi| e^{2\pi i \theta} : |\xi| \in [\frac{3}{4}\mathbf{ann}, \frac{7}{4}\mathbf{ann}), \theta \in \omega\}.$$

and

$$\mathbf{S}_{\mathbf{ann}, \omega} = \{R \times \Omega_{\mathbf{ann}, \omega} : R \in \mathcal{R}_{\mathbf{ann}, \omega}\}.$$

Here  $\mathcal{R}_{\mathbf{ann}, \omega}$  denotes all the rectangles with dimensions  $\mathbf{ann}^{-1}$  and  $|\omega|^{-1} |\mathbf{ann}|^{-1}$ , which are obtained by rotating with angle  $c(\omega)$  the rectangles with the same dimensions from  $\mathcal{D} \times \mathcal{D}$ . The elements in the collection

$$\mathbf{S}_u = \bigcup_{\omega, \mathbf{ann}} \mathbf{S}_{\mathbf{ann}, \omega}$$

will be referred to as *tiles*. We will typically write a tile  $s$  as  $s = R_s \times \Omega_s$ , and also denote by  $\mathbf{ann}(s)$ ,  $\omega(s)$  the corresponding components. Also,  $\omega_{1s}, \omega_{2s}$  will be the left and right dyadic children of  $\omega_s$ , while  $\Omega_{1s}, \Omega_{2s}$  will denote the corresponding sectors. The number  $\text{ecc}(s) := |\omega_s|$  will be referred to as the eccentricity of  $s$ .

Let  $\Phi$  be a positive Schwartz function supported on  $[\frac{3}{4}, \frac{7}{4}]$  with

$$\sum_{\mathbf{ann}} \Phi_{\mathbf{ann}} = 1,$$

where  $\Phi_{\mathbf{ann}}(t) = \Phi(\mathbf{ann}^{-1}t)$ .

The next step is the partition of unity for  $\mathbf{1}_\alpha$  described in Chapter 7 of [31]. Following this approach, the analysis of  $\widehat{G_\alpha f}$  is reduced to that of operators of the form

$$\sum_{\omega \subset [0, 1]: \omega \in \mathcal{G}} \hat{f}(\xi) \beta_\omega(\theta(\xi)) \mathbf{1}_{\omega_2}(\alpha).$$

Here  $\mathcal{G}$  is a fixed dyadic grid, independent of  $\alpha$ . For all practical purposes, we can assume  $\mathcal{G}$  is in fact the standard dyadic grid  $\mathcal{D}$ . We will denote by  $\omega_1, \omega_2$  the left and right dyadic children of  $\omega \in \mathcal{D}$ . Also, each  $\beta_\omega$  is a smooth bump function supported on  $\omega_1$  such that

$$(5.1) \quad \|\beta_\omega^{(j)}\|_\infty \leq C_j |\omega|^{-j}, \quad j \geq 0.$$

Finally, note that the multiplier

$$m_{\text{ann},\omega}(\xi) = \beta_\omega(\theta(\xi)) \Phi_{\text{ann}}(|\xi|)$$

is supported in and  $L^\infty$  adapted to the rectangle-like region  $\Omega_{\text{ann},\omega_1}$ . By applying a standard windowed Fourier series we can write

$$(5.2) \quad \hat{f}(\xi) m_{\text{ann},\omega}(\xi) = \sum_{s \in \mathbf{S}_{\text{ann},\omega}} \langle f, \varphi_s \rangle \hat{\psi}_s(\xi),$$

where for each  $s = R_s \times \Omega_s \in \mathbf{S}_{\text{ann},\omega}$ ,  $\varphi_s$  and  $\psi_s$  are Schwartz functions supported in frequency in  $\Omega_{\text{ann},\omega_{1s}}$  and  $L^2$ -adapted to the rectangle  $R_s$ .

By putting all these things together, it follows that we can write  $G_\alpha f(x)$  as a combination of model sums of the form

$$(5.3) \quad \sum_{s \in \mathbf{S}_u} \langle f, \varphi_s \rangle \psi_s(x) 1_{\omega_{2s}}(\alpha).$$

By splitting the model sum in two parts, we can further assume that  $\text{ann}(s) \in 4^\mathbb{Z}$  for each  $s \in \mathbf{S}_u$ . In particular, if  $\text{ann}(s) \neq \text{ann}(s')$  then  $\langle \varphi_s, \varphi_{s'} \rangle = \langle \psi_s, \psi_{s'} \rangle = 0$ .

For the purpose of future applications of the theory in Section 4, we remark that  $\varphi_s$  and  $\psi_s$  are in fact *product* wave packets adapted to  $R_s$ . This follows since the canonical directions  $e^{2\pi i c(\omega_s)}, e^{i\frac{\pi}{2} + 2\pi i c(\omega_s)}$  of  $R_s$  do not intersect the frequency support  $\Omega_{1s}$ . The same argument shows that  $\varphi_s$  and  $\psi_s$  continue to have the same property if  $R_s$  is slightly rotated so that the canonical directions become  $v, v^\perp$ , for any  $v \in \omega_{2s}$ .

Finally, we mention that the discretization (5.3) will serve us well when we treat the operator  $H_{\mathbf{V}}$  in the following sections. The fact that we have a sharp cutoff  $1_{\omega_{2s}}(\alpha)$  in the model sum will be absolutely crucial in our argument. We do not know if this decomposition can also be achieved for general kernels  $K$  (that is for the operators  $T_{\mathbf{V}}$ ). All earlier papers on the vector field problem used model sums involving smooth cutoffs  $\beta_{\omega_{2s}}(\alpha)$  ( $\beta$  smooth), which can in turn be realized as building blocks for all  $T_{\mathbf{V}}$ . The fact that our argument needs rough cutoffs is due to our use of product trees. In short, any smooth model for a product tree will produce error terms that are too large to be considered negligible.

## 6. MODEL SUMS, TREES, AND AN $L^p$ ALMOST ORTHOGONALITY PRINCIPLE

**6.1. Trees, size, and an almost orthogonality principle.** Let  $\mathbf{S}$  denote a fixed arbitrary finite subset of  $\mathbf{S}_u$ . We call *trees* those subsets of  $\mathbf{S}_u$  which point roughly in the same direction.

6.1.1. *Trees and size.* A subset  $\mathbf{T} \subset \mathbf{S}_u$  is a *lacunary tree* if

$$(6.1) \quad \omega_{2\mathbf{T}} = \bigcap_{s \in \mathbf{T}} \omega_{2s} \neq \emptyset.$$

Thus, for a lacunary tree  $\mathbf{T}$ , the Heisenberg boxes  $R_s \times \Omega_{1s}$  of  $\varphi_s$  are pairwise disjoint, and it is easy to see that we have the almost orthogonality relation

$$(6.2) \quad \sum_{s \in \mathbf{T}} |\langle f, \varphi_s \rangle|^2 \lesssim \|f\|_2^2.$$

A subset  $\mathbf{T} \subset \mathbf{S}_u$  is an *overlapping tree* if

$$(6.3) \quad \omega_{1\mathbf{T}} = \bigcap_{s \in \mathbf{T}} \omega_{1s} \neq \emptyset.$$

For an overlapping tree  $\mathbf{T}$ , the directional support uncertainty intervals  $\{\omega_{2s} : s \in \mathbf{T}\}$  are pairwise disjoint.

A subset  $\mathbf{t} \subset \mathbf{S}$  is a *conical tree* if  $\omega_{2s} = \omega_{2s'} =: \omega_{\mathbf{t}}$ , for all  $s, s' \in \mathbf{t}$ . A conical tree is both lacunary and overlapping. Any  $v_{\mathbf{T}} \in \omega_{2\mathbf{T}}$  (resp.  $v_{\mathbf{T}} \in \omega_{1\mathbf{T}}$ ,  $v_{\mathbf{t}} \in \omega_{\mathbf{t}}$ ) is a *top direction* of the lacunary (resp. overlapping, conical) tree  $\mathbf{T}$  (resp.  $\mathbf{t}$ ). If  $\mathbf{T}$  is a lacunary tree and  $s \in \mathbf{T}$ , then

$$(6.4) \quad \int_{\{x+tv_{\mathbf{T}}: t \in \mathbb{R}\}} \varphi_s = \int_{\{x+tv_{\mathbf{T}}: t \in \mathbb{R}\}} \varphi_s = 0, \quad \forall x \in \mathbb{R}^2.$$

The *shadow* and *k-shadow* of a tree are respectively defined as

$$\text{sh}(\mathbf{T}) = \cup \{R_s : s \in \mathbf{T}\}, \quad \text{sh}_k(\mathbf{T}) = \cup \{R_s^{(k)} : s \in \mathbf{T}\},$$

where  $R_s^{(k)} = 2^k R_s$ . For future use, we also introduce the notation

$$(6.5) \quad \text{crown}(\mathbf{T}) = \bigcup_{s \in \mathbf{T}} \omega_{2s} \subset (0, 1) \sim S^1$$

for the *crown* of a tree  $\mathbf{T}$ .

Let  $f$  be a given function. The *lacunary size* of  $\mathbf{S}' \subset \mathbf{S}$  (with respect to  $f$ , whose dependence in the notation is suppressed) is given by

$$(6.6) \quad \text{size}(\mathbf{S}') = \sup_{\substack{\mathbf{T} \subset \mathbf{S}' \\ \text{lacunary tree}}} \left( \frac{1}{|\text{sh}(\mathbf{T})|} \sum_{s \in \mathbf{T}} |\langle f, \varphi_s \rangle|^2 \right)$$

where  $\varphi_s$  are product wave packets adapted to  $\mathbf{S}$ . The *conical size* of  $\mathbf{S}' \subset \mathbf{S}$  is defined as

$$(6.7) \quad \text{size}_{\nabla}(\mathbf{S}') = \sup_{\substack{\mathbf{t} \subset \mathbf{S}' \\ \text{conical tree}}} \left( \frac{1}{|\text{sh}(\mathbf{t})|} \sum_{s \in \mathbf{t}} |\langle f, \varphi_s \rangle|^2 \right).$$

Since conical trees are in particular lacunary trees, we always have  $\text{size}_{\nabla}(\mathbf{S}') \leq \text{size}(\mathbf{S}')$ . We recall for future use [21, Lemma 4.40]

$$(6.8) \quad \left( \frac{1}{|\text{sh}(\mathbf{T})|} \sum_{s \in \mathbf{T}} |\langle f, \varphi_s \rangle|^2 \right)^{1/2} \lesssim \|f\|_{\infty},$$

for each lacunary  $\mathbf{T}$ .

6.1.2. *Single tree operators.* Let  $\mathbf{T}$  be a lacunary tree. Let  $\{\varphi_s, \psi_s\}_{s \in \mathbf{T}}$ , and for some  $k \geq 0$ ,  $\{\psi_s^{(k)}\}_{s \in \mathbf{T}}$ , be product wave packets adapted to  $\{R_s\}_{s \in \mathbf{T}}$  (resp.  $\{R_s^{(k)}\}_{s \in \mathbf{T}}$ ). Assume in addition that  $\text{supp } \psi_s^{(k)} \subset R_s^{(k)}$ . Define the model sums

$$(6.9) \quad f_{\mathbf{T}} = \sum_{s \in \mathbf{T}} \langle f, \varphi_s \rangle \psi_s,$$

and

$$(6.10) \quad f_{\mathbf{T}}^{(k)} = \sum_{s \in \mathbf{T}} \langle f, \varphi_s \rangle \psi_s^{(k)}.$$

**Lemma 6.1.** *We have the estimate*

$$(6.11) \quad \|f_{\mathbf{T}}\|_p \leq Cp^2 \text{size}(\mathbf{T}), \quad 2 \leq p < \infty.$$

**Lemma 6.2.** *We have the estimate*

$$(6.12) \quad \left| \{x \in \text{sh}_k(\mathbf{T}) : |f_{\mathbf{T}}^{(k)}(x)| > \lambda\} \right| \leq C \exp \left( -c \sqrt{\frac{\lambda}{\text{size}(\mathbf{T})}} \right) |\text{sh}_k(\mathbf{T})|.$$

*Proofs of Lemmata 6.1 and 6.2.* Estimates (6.11) and (6.12) follow respectively from (4.28) and (4.10). Let us explain the details for (6.11). Note that for each  $s \in \mathbf{S}$ ,  $v_{\mathbf{T}} \in \omega_{2s}$ , which is within the uncertainty angle of  $R_s$ . Therefore, property (4.5) still holds if we replace  $e_{R_s}$  by  $v_{\mathbf{T}}$ . Then, we can find a rectangle  $\rho_s$  oriented along  $v_{\mathbf{T}}$  such that  $R_s \subset \rho_s \subset 90R_s$  and  $\{\rho_s : s \in \mathbf{T}\}$  are products of intervals coming from finitely many dyadic grids. Thus one observes that setting  $b_{\rho_s} := \langle f, \varphi_s \rangle$ , gives  $f_{\mathbf{T}} = \mathbf{B}_{\{\rho_s : s \in \mathbf{T}\}}$ , and  $\{\varphi_s, \psi_s\}$  are still product wave packets adapted to  $\{\rho_s : s \in \mathbf{T}\}$ . Therefore (4.28) applies.  $\square$

An application of (4.11) yields a similar result for conical trees. Let  $\mathbf{t}$  be a conical tree. Let  $\{\varphi_s\}_{s \in \mathbf{t}}$ , and for some  $k \geq 0$ ,  $\{\psi_s^{(k)}\}_{s \in \mathbf{t}}$ , be wave packets adapted to  $\{R_s\}_{s \in \mathbf{t}}$  (resp.  $\{R_s^{(k)}\}_{s \in \mathbf{t}}$ ). Then

$$(6.13) \quad \left| \{x \in \text{sh}_k(\mathbf{t}) : |f_{\mathbf{t}}^{(k)}(x)| > \lambda\} \right| \leq C \exp \left( -c \frac{\lambda}{\text{size}_{\nabla}(B)} \right) |\text{sh}_k(\mathbf{t})|.$$

6.1.3. *An  $L^p$  almost orthogonality principle and an application to  $H_{\mathbf{V}}$ .* We prove the following  $L^p$  almost orthogonality result.

**Theorem 4.** *Let  $\alpha$  be a collection of disjoint intervals. We have the estimate*

$$(6.14) \quad \left( \sum_{\alpha \in \alpha} \|G_{\alpha} f\|_p^p \right)^{\frac{1}{p}} \lesssim_p \|f\|_p, \quad \forall p \geq 2.$$

*Proof.* By restricted-type interpolation, it suffices to show that (6.14) holds for  $|f| \leq \mathbf{1}_E$ , where  $E \subset \mathbb{R}^2$  has finite measure. We want to use a model sum similar to the one in (5.3). However, we use a more straightforward partition of unity for  $\alpha = (\alpha_l, \alpha_r)$ , namely

$$1_{\alpha} = \sum_{\omega \in C_{\alpha}} \beta_{\omega}.$$

Here  $C_{\alpha}$  consists of all intervals  $\omega \in \mathcal{D} \cup \mathcal{D}^1 \cup \mathcal{D}^2$  such that  $\omega \subset \alpha$  and  $\text{dist}(\omega, \{\alpha_l, \alpha_r\}) \sim |\omega|$ , while  $\beta_{\omega}$  is supported in and  $L^{\infty}$ -adapted to  $\omega$ . Since  $\|\sum_{\omega \in C_{\alpha}} 1_{\omega}\|_{\infty} \lesssim 1$ , by working



with a finite number of subsets we can assume the intervals in  $C_\alpha$  are pairwise disjoint,  $\text{dist}(\omega, \alpha_l) \sim |\omega|$  and prove

$$(6.15) \quad \left( \sum_{\alpha \in \mathbf{\alpha}} \|f_{\mathbf{T}(\alpha)}\|_p^p \right)^{\frac{1}{p}} \lesssim_p \|f\|_p, \quad \forall p > 2.$$

Here

$$\mathbf{T}(\alpha) = \{s \in \mathbf{S}_u : \omega_s \in C_\alpha\},$$

$$f_{\mathbf{T}(\alpha)} = \sum_{\omega_s \in \mathbf{T}(\alpha)} \langle f, \varphi_s \rangle \psi_s,$$

and for each  $s = R_s \times \Omega_s$ ,  $\varphi_s, \psi_s$  are product wave packets adapted to  $R_s$  and supported in frequency in  $\Omega_s$ . The axes of  $R_s$  are oriented along  $e^{2\pi i \alpha_l}, e^{\frac{\pi}{2} + 2\pi i \alpha_l}$

Note that  $\mathbf{T}(\alpha)$  is similar in nature to a lacunary tree, in particular (6.2), (6.8) and (6.11) will also hold for it. We chop each  $\mathbf{T}(\alpha)$  into subsets  $\mathbf{T}_\sigma(\alpha)$  by means of the following standard iterative procedure.

- INIT Stock :=  $\mathbf{T}(\alpha)$ ,  $\sigma \sim 1$ ;
- WHILE Stock  $\neq \emptyset$ , select the largest collection  $\mathbf{T}_\sigma(\alpha) \subset \text{Stock}$  with

$$\sum_{s \in \mathbf{T}_\sigma(\alpha)} |\langle f, \varphi_s \rangle|^2 \geq \frac{\sigma^2}{4} |\text{sh}(\mathbf{T}_\sigma(\alpha))|.$$

$$\text{Set Stock} := \text{Stock} \setminus \mathbf{T}_\sigma(\alpha); \sigma := \frac{\sigma}{2}.$$

This iterative process produces the decomposition

$$(6.16) \quad \mathbf{T}(\alpha) = \bigcup_{\sigma \lesssim 1} \mathbf{T}_\sigma(\alpha), \quad \text{size}(\mathbf{T}_\sigma(\alpha)) \leq \sigma,$$

and the almost-orthogonality (6.2) of the  $\{\varphi_s : s \in \mathbf{T}(\alpha)\}$  yields

$$(6.17) \quad |\text{sh}(\mathbf{T}_\sigma(\alpha))| \lesssim \sigma^{-2} \sum_{s \in \mathbf{T}_\sigma(\alpha)} |\langle f, \varphi_s \rangle|^2 \lesssim \sigma^{-2} \|G_\alpha f\|_2^2.$$

Therefore

$$\begin{aligned} \left( \sum_{\alpha \in \mathbf{\alpha}} \|f_{\mathbf{T}(\alpha)}\|_p^p \right)^{\frac{1}{p}} &= \left\| \|f_{\mathbf{T}(\alpha)}\|_{\ell^p(\mathbf{\alpha})} \right\|_p = \left\| \left\| \sum_{\sigma \lesssim 1} f_{\mathbf{T}_\sigma(\alpha)} \right\|_{\ell^p(\mathbf{\omega})} \right\|_p \\ &\leq \sum_{\sigma \lesssim 1} \left( \sum_{\alpha \in \mathbf{\alpha}} \|f_{\mathbf{T}_\sigma(\alpha)}\|_p^p \right)^{\frac{1}{p}} \lesssim \sum_{\sigma \lesssim 1} \sigma \left( \sum_{\alpha \in \mathbf{\alpha}} |\text{sh}(\mathbf{T}_\sigma(\alpha))| \right)^{\frac{1}{p}} \\ &\lesssim \sum_{\sigma \lesssim 1} \sigma^{1-\frac{2}{p}} \left( \sum_{\alpha \in \mathbf{\alpha}} \|G_\alpha f\|_2^2 \right)^{\frac{1}{p}} \lesssim (\|f\|_2^2)^{\frac{1}{p}} = |E|^{\frac{1}{p}}. \end{aligned}$$

In the above display, we used the single tree estimate (6.11) of Lemma 6.1 to get the second inequality of the second line, (6.17) in going from the second to the third line, and disjointness of  $\alpha \in \mathbf{\alpha}$  to conclude. This proves (6.15), and in turn, Theorem 4.  $\square$

We can use the  $L^p$ -orthogonality principle of Theorem 4 to give a simple proof of the bound

$$\|H_{\mathbf{V}}f\|_p \lesssim_p \log N \|f\|_p, \quad 2 < p < \infty.$$

The proof is done by induction on  $n$ , where  $N = 2^n$ . For the inductive step, assume that whenever  $\#\mathbf{V}' \leq 2^{n-1}$ ,  $\|H_{\mathbf{V}'}\|_{p \rightarrow p} \leq C_p(n-1)$ , where  $C_p$  is bigger than the implicit constant appearing in Theorem 4. Order  $\mathbf{V}$  increasingly and let  $\mathbf{V}' = \{v_i \in \mathbf{V} : i \text{ odd}\}$ . For  $i$  odd, let  $\omega_i$  be the frequency cone  $\{\xi : \frac{\xi}{|\xi|} \in \frac{\pi}{2} + [v_i, v_{i+2})\}$ . For each  $i$  odd,

$$(6.18) \quad v \in \{v_i, v_{i+1}\} \implies |H_v f(x)| \leq |H_{v_i} f(x)| + |H_{v_{i+1}}[G_{\omega_i} f](x)|,$$

and taking suprema and  $L^p$  norms,

$$\begin{aligned} \|H_{\mathbf{V}}f\|_p &\leq \|H_{\mathbf{V}'}f\|_p + \left\| \sup_{i \text{ odd}} H_{v_{i+1}}[G_{\omega_i} f] \right\|_p \\ &\leq C_p(n-1) \|f\|_p + \left( \sum_i \|G_{\omega_i} f\|_p^p \right)^{\frac{1}{p}} \\ &\leq C_p n \|f\|_p, \end{aligned}$$

which completes the induction. The last line is an application of Theorem 4, while, in going from the first to the second line, we have used the induction hypothesis and the fact that only one  $i$  contributes at each point.

**6.2. The model sum for  $H_{\mathbf{V}}$ .** We recall that, for a direction  $v$ , the directional multiplier  $H_v$  is given by  $\widehat{H_v f}(\xi) = \hat{f}(\xi) \text{sign}(v \cdot \xi)$ , which is essentially the same as

$$\hat{f}(\xi) \mathbf{1}_{(0, \tau(v))}(\theta), \quad \xi = |\xi| e^{2\pi i \theta},$$

where  $v = \exp(2\pi i \tau(v) - \frac{\pi i}{2})$ . Then, we use (5.3) to define the model sum for  $H_v$

$$(6.19) \quad H_{\mathbf{S}} f(x, v) = \sum_{s \in \mathbf{S}} \langle f, \varphi_s \rangle \psi_s(x) \mathbf{1}_{\omega_{2s}}(v),$$

where  $\mathbf{S}$  is any finite subcollection of  $\mathbf{S}_u$ , and we identify  $\tau(v)$  with  $v$ . Accordingly, the model operator for the maximal multiplier  $H_{\mathbf{V}}$  is defined as

$$(6.20) \quad H_{\mathbf{S}}^* f(x) = \sup_{v \in \mathbf{V}} |H_{\mathbf{S}} f(x, v)|.$$

Finally, we write  $H_{\mathbf{S}}^*$  as a rapidly decaying sum of pieces supported in the dilates  $\text{sh}_k(\mathbf{S})$ ,  $k \geq 0$ . Applying Lemma 4.6 to each  $\psi_s$ , we write  $\psi_s = \sum_{k \geq 0} 2^{-100k} \psi_s^{(k)}$ , where  $\psi_s^{(k)}$  are product wave packets adapted to the rectangles  $\{R_s^{(k)} := 2^k R_s\}$ , with the mean zero property

$$(6.21) \quad v \in \omega_{2s} \implies \int_{\{x+tv: t \in \mathbb{R}\}} \psi_s^{(k)} = \int_{\{x+tv^\perp: t \in \mathbb{R}\}} \psi_s^{(k)} = 0, \quad \forall x \in \mathbb{R}^2.$$

We thus have

$$H_{\mathbf{S}} f(x, v) = \sum_{k \geq 0} 2^{-100k} \sum_{s \in \mathbf{S}} \langle f, \varphi_s \rangle \psi_s^{(k)}(x) \mathbf{1}_{v \in \omega_{2s}}$$

Defining

$$H_{\mathbf{S}(k)}^* f(x) = \sup_{v \in \mathbf{V}} \left| \sum_{s \in \mathbf{S}} \langle f, \varphi_s \rangle \psi_s^{(k)}(x) \mathbf{1}_{v \in \omega_{2s}} \right|,$$

we have the pointwise inequality

$$|H_{\mathbf{S}}^* f(x)| \lesssim \sum_{k \geq 0} 2^{-100k} |H_{\mathbf{S},k}^* f(x)|.$$

In particular, we reduced Theorem 3 to the proof of the weak-type bound

$$(6.22) \quad \|H_{\mathbf{S}(k)}^* f\|_{2,\infty} \lesssim_Q 2^{4k} \sqrt{\log N} (\log \log N)^6.$$

## 7. VARGAS TYPE $\mathbf{V}$ : PROOF OF THEOREM 3

Before entering the proof, we begin the section with a more detailed description of Vargas sets of directions, and provide some important examples.

**7.1. Vargas sets.** We recall Definition 1.3. We say that a set of  $N$  directions is a Vargas set with constant  $Q$  if the longest lacunary sequence (see Definition 1.2) in  $\mathbf{V}$  has  $Q \log N$  elements. Our definition is easily seen to be equivalent to the one of [18], where the terminology has been introduced. For example, uniformly distributed sets

$$\mathbf{V}_N = \{v_j = \frac{j}{2\pi N}, j = 0, \dots, N-1\}$$

are of (uniform in  $N$ ) Vargas sets with constant 1. Another prime example of (uniformly) Vargas sets are finite truncations of Cantor-type sets.

CLAIM. The truncated  $q$ -adic ( $q \geq 3$ ) Cantor sets of  $N = 2^n$  elements

$$\mathbf{V}_N := \mathfrak{C}_{q,n} = \left\{ \sum_{j=1}^n a_j q^{-j}, (a_1, \dots, a_n) \in \{0, q-1\}^n \right\},$$

are Vargas sets with constant 4 (independent of  $q$  and  $N$ ).

*Proof.* We argue by induction on  $n$ . The case  $n = 1$  is trivial. The inductive assumption is that for each  $\nu < n$ , a lacunary sequence contained in  $\mathfrak{C}_{q,\nu}$  has at most  $4\nu$  elements. Take a lacunary sequence

$$\left\{ v_\ell = \sum_{j=1}^n v_\ell^{(j)} q^{-j}, \ell = 1, \dots, r \right\} \subset \mathfrak{C}_{q,n}$$

Let  $h$  be the first integer such that  $v_1^{(h)} \neq v_2^{(h)}$ . Then

$$\frac{q-1}{q^h} \leq |v_1 - v_2| \leq \frac{1}{q^{h-1}}.$$

As a consequence of the lacunarity property (1.9), it is easy to see that

$$|v_\ell - v_m| \leq \frac{3}{2^{m-2}} |v_2 - v_1|, \quad \forall m \geq 2, \ell = m, \dots, r.$$

Hence, for each  $\ell = 5, \dots, r$ , we must have  $v_\ell^{(j)} = v_5^{(j)}$ , for each  $j = 1, \dots, h$ . If not, there would be some  $j \leq h, \ell \geq 5$  such that

$$v_\ell^{(j)} \neq v_5^{(j)} \implies \frac{q-1}{q^j} \leq |v_\ell - v_5| \leq \frac{3}{8} |v_1 - v_2| \leq \frac{3}{8} \frac{1}{q^{h-1}},$$

which gives the contradiction  $q - 1 \leq \frac{3}{8}q$ . This shows that  $\{v_5, \dots, v_r\}$  is a lacunary sequence contained in

$$\left\{ \sum_{j=1}^n a^{(j)} q^{-j}, a^{(1)} = v_5^{(1)}, \dots, a^{(h)} = v_5^{(h)}, (a^{(h+1)}, \dots, a^{(n)}) \in \{0, q-1\}^{n-h} \right\} \sim \mathfrak{C}_{q, n-h},$$

so that  $r - 4 \leq 4(n - h)$  by the inductive assumption. We get  $r \leq 4n - 4h + 4 \leq 4n$ , which completes the inductive step.  $\square$

The claim below shows how Vargas sets and lacunary sets are, in a sense, extremes.

CLAIM. Let  $\#\mathbf{V} = N$ . Then  $\mathbf{V}$  contains a lacunary sequence of  $\frac{\log N}{3}$  elements.

*Proof.* This is best seen by induction on  $n = \frac{\log N}{3}$ . Take as base case  $n = 4$ . The inductive assumption is that every set of at least  $2^{3\nu}$  directions contains a lacunary sequence of  $\nu$  elements, for  $4 \leq \nu \leq n$ . Let  $\mathbf{V} = \{v_1 < v_2 < \dots v_{2^{3(n+1)}}\}$  be given. We can rescale and assume that  $v_1 = 0$ ,  $v_{2^{3(n+1)}} = 1$ . Define  $\mathbf{V}_j = \mathbf{V} \cap [\frac{j-1}{2^3}, \frac{j}{2^3}]$ ,  $j = 1, \dots, 8$ . By pigeonholing, at least one  $\mathbf{V}_j$  has at least  $2^{3n}$  elements: assume it is so for  $\mathbf{V}_2$  (the other cases are treated similarly). Then, by inductive assumption,  $\mathbf{V}_2$  contains a lacunary sequence  $\{w_1, \dots, w_n\} \subset [\frac{1}{8}, \frac{1}{4}]$  with node  $w$ . Moreover, by repeated application of (1.9),  $\text{dist}(w, \mathbf{V}_2) \leq 2^{-(n-1)}$  so that

$$(7.1) \quad |w - w_1| \leq \text{diam}(\mathbf{V}_2) + \text{dist}(w, \mathbf{V}_2) \leq \frac{1}{8} + 2^{-(n-1)} \leq \frac{1}{4}.$$

We set  $w_0 := v_{2^{3(n+1)}} = 1$ . Then  $w \in [0, \frac{1}{2}]$ , so that  $|w - w_{n+1}| \geq \frac{1}{2}$ . Comparing this to (7.1), we observe that  $\{w_0, w_1, \dots, w_n\}$  is a lacunary sequence of  $n+1$  elements contained in  $\mathbf{V}$ . The inductive step is complete.  $\square$

**7.2. Preliminary reductions and outline of the proof.** Let  $\mathbf{V}$  be a fixed set of directions with  $\#\mathbf{V} = N$  and Vargas constant  $Q$ . Using the notations and the results of Section 7, Theorem 3 follows from the weak-type bound

$$\|\mathbf{H}_{\mathbf{S}^{(k)}}^* f\|_{2,\infty} \lesssim_Q 2^{4k} \sqrt{\log N} (\log \log N)^6.$$

From now on, the dependence on  $Q$  of the estimates is kept implicit. By scaling  $f$ , it will suffice to show that

$$(7.2) \quad |\{x \in \mathbb{R}^2 : \mathbf{H}_{\mathbf{S},k}^* f(x) \gtrsim 2^{2k} \sqrt{\log N} (\log \log N)^6\}| \lesssim 2^{4k} \|f\|_2^2.$$

We will obtain (7.2) by constructing a set  $A = A^1 \cup A^2 \subset \mathbb{R}^2$  with the properties

$$(7.3) \quad |A| \lesssim 2^{4k} \|f\|_2^2, \quad \mathbf{H}_{\mathbf{S}^{(k)}}^* f(x) \lesssim 2^{2k} \sqrt{\log N} (\log \log N)^6 \quad \forall x \notin A.$$

This will be done by decomposing  $\mathbf{S}$  into the disjoint union of forests (i.e. collections of trees)  $\mathbf{S}_\sigma$ , where  $\sigma$  is a dyadic parameter ranging in  $\{z \in 2^\mathbb{Z} : z \leq 2^K\}$ , which stands for the size of the subcollection  $\mathbf{S}_\sigma$ , as we will detail later. Here,  $K = \text{size}(\mathbf{S})$ . Accordingly, using the triangle inequality

$$(7.4) \quad \mathbf{H}_{\mathbf{S}^{(k)}}^* f(x) \leq \sum_{\sigma \leq 2^K} \mathbf{H}_{\mathbf{S}_\sigma^{(k)}}^* f(x),$$

and each piece  $H_{\mathbf{S}_\sigma}^* f$  is supported on  $\text{sh}_k(\mathbf{S}_\sigma)$ . In turn, each  $\mathbf{S}_\sigma$  will be the union of trees  $\mathbf{T} \in \mathcal{F}_\sigma$  with the properties

$$(7.5) \quad \sum_{\mathbf{T} \in \mathcal{F}_\sigma} |\text{sh}(\mathbf{T})| \lesssim \sigma^{-2} \|f\|_2^2,$$

$$(7.6) \quad \sup_{v \in \mathbf{V}} \sum_{\mathbf{T} \in \mathcal{F}_\sigma} \mathbf{1}_{\text{crown}(\mathbf{T})}(v) \leq C \log N.$$

Property (7.6) will be obtained as a consequence of the Vargas set structure of  $\mathbf{V}$ . We take the first component of the exceptional set  $A$  to be

$$A^1 = \bigcup_{1 \leq \sigma \leq 2^K} \bigcup_{\mathbf{T} \in \mathcal{F}_\sigma} \text{sh}_k(\mathbf{T}).$$

Using (7.5), we have

$$|A^1| \leq \sum_{1 \leq \sigma \leq 2^K} \sum_{\mathbf{T} \in \mathcal{F}_\sigma} |\text{sh}_k(\mathbf{T})| \lesssim 2^{4k} \sum_{1 \leq \sigma \leq 2^K} \sigma^{-2} \|f\|_2^2 \lesssim 2^{4k} \|f\|_2^2.$$

The support of each  $H_{\mathbf{S}_\sigma}^*$ ,  $1 \leq \sigma \leq 2^K$  is contained in  $A^1$ , so that

$$(7.7) \quad H_{\mathbf{S}_\sigma}^* f(x) \leq \sum_{\sigma \leq 1} H_{\mathbf{S}_\sigma}^* f(x) \leq \sum_{\sigma \leq 1} \sum_{\mathbf{T} \in \mathcal{F}_\sigma} H_{\mathbf{T}(k)} f(x), \quad x \notin A^1.$$

We now construct the second component of  $A$ . We begin by excising an exceptional set for each tree  $\mathbf{T} \in \mathcal{F}_\sigma$  appearing in the sum (7.7): setting

$$(7.8) \quad A_{\mathbf{T}, \sigma}^{2, \text{size}} = \{x \in \text{sh}_k(\mathbf{T}) : H_{\mathbf{T}(k)}^* f(x) \gtrsim 2^{2k} \sigma |10 \log \sigma|^2 (\log \log N)^2\}, \quad \mathbf{T} \in \mathcal{F}_\sigma$$

we will show that

$$(7.9) \quad |A_{\mathbf{T}, \sigma}^{2, \text{size}}| \lesssim \sigma^{10} 2^{2k} |\text{sh}(\mathbf{T})|.$$

Set  $A_\sigma^{2, \text{size}} := \bigcup_{\mathbf{T} \in \mathcal{F}_\sigma} A_{\mathbf{T}, \sigma}^{2, \text{size}}$ . Using (7.5), we have the estimate

$$(7.10) \quad |A_\sigma^{2, \text{size}}| \leq \sum_{\mathbf{T} \in \mathcal{F}_\sigma} |A_{\mathbf{T}, \sigma}^{2, \text{size}}| \leq 2^{2k} \sigma^{10} \sum_{\mathbf{T} \in \mathcal{F}_\sigma} |\text{sh}(\mathbf{T})| \lesssim 2^{2k} \sigma^{-8} \|f\|_2^2$$

We also excise an exceptional set coming from the whole forest  $\mathbf{S}_\sigma$ :

$$(7.11) \quad A_\sigma^{2, \text{count}} := \left\{ \sum_{\mathbf{T} \in \mathcal{F}_\sigma} \mathbf{1}_{\text{sh}_k(\mathbf{T})}(x) \gtrsim \sigma^{-2} |\log \sigma|^2 \right\}.$$

Another application of (7.5) entails

$$(7.12) \quad |A_\sigma^{2, \text{count}}| \leq \frac{\sigma^{-2}}{|\log \sigma|^2} \sum_{\mathbf{T} \in \mathcal{F}_\sigma} |\text{sh}_k(\mathbf{T})| \lesssim \frac{2^{4k}}{|\log \sigma|^2} \|f\|_2^2.$$

Putting all pieces together, we obtain our second component  $A^2$ :

$$(7.13) \quad A^2 = \bigcup_{\sigma \lesssim 1} (A_\sigma^{2, \text{size}} \cup A_\sigma^{2, \text{count}}), \quad |A^2| \lesssim 2^{4k} \|f\|_2^2.$$

the estimate on the measure of  $A^2$  is obtained by summing over  $\sigma \lesssim 1$  the estimates in (7.10) and (7.12). Thus, the first part of (7.3) holds true.

We are left to show the second part of (7.3). Recall that  $\mathbf{crown}(\mathbf{T}) = \bigcup\{\omega = \omega_{2s} : s \in \mathbf{T}\}$ . As a consequence of the definition of  $A_\sigma^{2,\text{count}}$  and of property (7.6), we have that

$$(7.14) \quad \sup_{x \notin A} \sup_{v \in \mathbf{V}} \sum_{\mathbf{T} \in \mathcal{F}_\sigma} \mathbf{1}_{\text{sh}_k(\mathbf{T})}(x) \mathbf{1}_{\mathbf{crown}(\mathbf{T})}(v) \lesssim \min\{\log N, \sigma^{-2} |\log \sigma|^2\}.$$

Then, for  $x \notin A$ , choose  $v(x) \in \mathbf{V}$  which attains the supremum in  $\mathbf{H}_{\mathbf{S}_\sigma}^*$ . We have

$$\begin{aligned} \mathbf{H}_{\mathbf{S}_\sigma}^* f(x) &= \left| \sum_{\mathbf{T} \in \mathcal{F}_\sigma} \sum_{s \in \mathbf{T}} \langle f, \varphi_s \rangle \psi_s^{(k)}(x) \mathbf{1}_{\omega_{2s}}(v(x)) \right| \\ &= \left| \sum_{\mathbf{T} \in \mathcal{F}_\sigma} \mathbf{1}_{\text{sh}_k(\mathbf{T})}(x) \mathbf{1}_{\mathbf{crown}(\mathbf{T})}(v(x)) \sum_{s \in \mathbf{T}} \langle f, \varphi_s \rangle \psi_s^{(k)}(x) \mathbf{1}_{v(x) \in \omega_{2s}} \right| \\ &\lesssim \min\{\log N, \sigma^{-2} |\log \sigma|^2\} \left( \sup_{\mathbf{T} \in \mathcal{F}_\sigma} \mathbf{H}_{\mathbf{T}^{(k)}}^* f(x) \right) \\ &\lesssim 2^{2k} (\log \log N)^2 \min\{\log N \sigma |\log \sigma|^2, \sigma^{-1} |\log \sigma|^4\} \end{aligned}$$

Taking advantage of (7.7), we then get, for  $x \notin A$ ,

$$\begin{aligned} \mathbf{H}_{\mathbf{S}_{(k)}}^* f(x) &\leq \sum_{\sigma \lesssim 1} \mathbf{H}_{\mathbf{S}_\sigma}^* f(x) \\ &\lesssim 2^{2k} (\log \log N)^2 \sum_{\sqrt{\log N} \geq \sigma^{-1} \geq 1} \sigma^{-1} |\log \sigma|^4 + 2^{2k} \log N (\log \log N)^2 \sum_{\sigma \lesssim \sqrt{\log N}} \sigma |\log \sigma|^2 \\ &\lesssim 2^{2k} \sqrt{\log N} (\log \log N)^6, \end{aligned}$$

which completes the proof of the second part of (7.3). This in turn proves the weak-type bound (7.2), and finally, Theorem 3. In the remainder of the section, we will show how to achieve a decomposition  $\mathbf{S} = \bigcup_{\sigma \leq 2^K} \mathbf{S}_\sigma$  complying with properties (7.5), (7.6), and whose trees satisfy the exceptional set estimate (7.9).

**7.3. Lacunary and overlapping parts of  $\mathbf{S}$ .** First of all, it suffices to assume that  $\#(\mathbf{V} \cap \omega_{2s}) \geq 1$  for each  $s \in \mathbf{S}$ . Indeed, tiles with empty  $\mathbf{V} \cap \omega_{2s}$  do not contribute to the model sum  $\mathbf{H}_{\mathbf{S}^{(k)}}^*$ . We decompose

$$\mathbf{S} = \mathbf{L} \cup \mathbf{O}, \quad \mathbf{L} = \{s \in \mathbf{S} : \mathbf{V} \cap \omega_{1s} = \emptyset\}, \quad \mathbf{O} = \{s \in \mathbf{S} : \mathbf{V} \cap \omega_{1s} \neq \emptyset\}$$

Then,  $\mathbf{S}_\sigma$  will be obtained as the union  $\mathbf{L}_\sigma \cup \mathbf{O}_\sigma$ . The forest  $\mathbf{L}_\sigma$  will be the union of lacunary trees  $\mathbf{T} \in \mathcal{F}_\sigma^{\mathbf{L}}$ , while the forest  $\mathbf{O}_\sigma$  will be the union of overlapping trees  $\mathbf{T} \in \mathcal{F}_\sigma^{\mathbf{O}}$ . For  $* \in \{\mathbf{L}, \mathbf{O}\}$ , and  $\sigma \leq 2^K$ , we will achieve the properties

$$(7.15) \quad \sum_{\mathbf{T} \in \mathcal{F}_\sigma^*} |\text{sh}(\mathbf{T})| \lesssim \sigma^{-2} \|f\|_2^2,$$

$$(7.16) \quad \sup_{v \in \mathbf{V}} \sum_{\mathbf{T} \in \mathcal{F}_\sigma^*} \mathbf{1}_{\mathbf{crown}(\mathbf{T})}(v) \leq C \log N,$$

$$(7.17) \quad \mathbf{T} \in \mathcal{F}_\sigma^* \implies |\{x : \mathbf{H}_{\mathbf{T}^{(k)}}^* f(x) \gtrsim 2^{2k} \sigma |\log \sigma|^2 (\log \log N)^2\}| \lesssim \sigma^{10} 2^{2k} |\text{sh}(\mathbf{T})|.$$

This will imply the corresponding properties for  $\mathbf{S}_\sigma$ , and in turn, complete the proof of Theorem 3.

**7.4. The lacunary part  $\mathbf{L}$ .** We begin by noting that the intervals  $\{\omega = \omega_{1s} : s \in \mathbf{L}\}$  are pairwise disjoint. Indeed, if we had  $\omega_{1s}$  strictly contained in  $\omega_{1s'}$ , for  $s, s' \in \mathbf{L}$ , then  $\omega_{2s} \subset \omega_{1s'}$  too. This would force  $\#(\mathbf{V} \cap \omega_{1s'}) \geq 1$ , which we ruled out. As a consequence, all trees  $\mathbf{T} \subset \mathbf{L}$  are of lacunary type, and the almost-orthogonality relation

$$(7.18) \quad \sum_{s \in \mathbf{L}} |\langle f, \varphi_s \rangle|^2 \lesssim \|f\|_2^2$$

holds. We organize  $\mathbf{L}$  into subcollections  $\mathbf{L}(j)$  by setting

$$\mathbf{L}(j) = \{s \in \mathbf{L} : 2^j \leq \omega_{2s} < 2^{j+1}\}, \quad j = 1, \dots, \log N.$$

Recall that  $\pi_s = \omega_{1s} \cup \omega_{2s}$ . For each  $\mathbf{L}(j)$ , we distinguish the minimal (with respect to inclusion) dyadic intervals  $\{\pi_\ell^j\}$  inside  $\{\pi = \pi_s : s \in \mathbf{L}(j)\}$ , and form the lacunary trees

$$\mathbf{T}(j, \ell) = \{s \in \mathbf{L}(j) : \omega_{2s} \supset \pi_\ell^j\}, \quad \mathbf{L}(j) = \cup_\ell \mathbf{T}(j, \ell).$$

It is clear that each  $\pi_s$  contains exactly one  $\pi_\ell^j$ . This forces  $\pi_s \cap \pi_{s'} = \emptyset$  whenever  $s \in \mathbf{T}(j, \ell), s' \in \mathbf{T}(j, \ell'), \ell' \neq \ell$ . Thus for each  $j$  the sets  $\{\text{crown}(\mathbf{T}(j, \ell)) : \ell\}$  are pairwise disjoint. We conclude that

$$(7.19) \quad \sup_{v \in S^1} \sum_{j=1}^{\log N} \sum_{\ell} \mathbf{1}_{\text{crown}(\mathbf{T}(j, \ell))}(v) \leq \log N.$$

We now construct  $\mathcal{F}_\sigma^\mathbf{L} = \cup_{j=1}^{\log N} \cup_\ell \mathbf{T}_\sigma(j, \ell)$  by chopping each tree  $\mathbf{T}(j, \ell)$  into subtrees with  $\text{size}(\mathbf{T}_\sigma(j, \ell)) \leq \sigma$ . The chopping is obtained via the following iterative procedure.

- **INIT**  $\text{Stock} := \mathbf{T}(j, \ell), \sigma = 2^K$ ;
- **WHILE**  $\text{Stock} \neq \emptyset$ , select the largest collection  $\mathbf{T}_\sigma(j, \ell) \subset \text{Stock}$  with

$$\sum_{s \in \mathbf{T}_\sigma(j, \ell)} |\langle f, \varphi_s \rangle|^2 \geq \frac{\sigma^2}{4} |\text{sh}(\mathbf{T}_\sigma(j, \ell))|.$$

Set  $\text{Stock} := \text{Stock} \setminus \mathbf{T}_\sigma(j, \ell); \sigma := \frac{\sigma}{2}$ .

The almost-orthogonality property (7.18) yields that

$$(7.20) \quad \sum_{j=1}^{\log N} \sum_{\ell} |\text{sh}(\mathbf{T}_\sigma(j, \ell))| \lesssim \sigma^{-2} \sum_{s \in \mathbf{L}} |\langle f, \varphi_s \rangle|^2 \lesssim \sigma^{-2} \|f\|_2^2,$$

so that (7.15) for  $* = \mathbf{L}$  is verified. Moreover, (7.16) follows *a fortiori* from (7.19). All that is left to verify is the single tree estimate (7.17). Let  $\mathbf{T} = \mathbf{T}(j, \ell) \in \mathcal{F}_\sigma^\mathbf{L}$  ( $j$  and  $\ell$  have no particular relevance). For each  $v \in \mathbf{V} \cap \text{crown}(\mathbf{T})$ , let us isolate the subtree of  $\mathbf{T}$  contributing at  $v$ , that is  $\mathbf{T}(v) = \{s \in \mathbf{T} : v \in \omega_{2s}\}$ . Set  $\mathcal{T}(\mathbf{T}) = \{\mathbf{T}(v) : v \in \mathbf{V} \cap \text{crown}(\mathbf{T})\}$ . With this position, using the notation of (6.10), we have

$$\mathbf{H}_{\mathbf{T}(k)}^* f(x) = \sup_{\mathbf{T}' \in \mathcal{T}(\mathbf{T})} |f_{\mathbf{T}'}^{(k)}(x)|.$$

We claim that, as a consequence of the Vargas set property,  $\mathcal{T}(\mathbf{T})$  contains at most  $C \log N$  distinct subtrees  $\mathbf{T}'$ . The proof of the claim is postponed at the end of the



section. Assuming for now that  $\#\mathcal{T}(\mathbf{T}) \leq C \log N$ , we can complete the proof of the single tree estimate as follows:

$$\begin{aligned}
(7.21) \quad & \left| \{x : \mathbf{H}_{\mathbf{T}^{(k)}}^* f(x) \gtrsim 2^{2k} \sigma |10 \log \sigma|^2 (\log \log N)^2\} \right| \\
& \leq \sum_{\mathbf{T}' \in \mathcal{T}(\mathbf{T})} \left| \{x : |f_{\mathbf{T}'}^{(k)}(x)| \gtrsim 2^{2k} \sigma |10 \log \sigma|^2 (\log \log N)^2\} \right| \\
& \leq \sum_{\mathbf{T}' \in \mathcal{T}(\mathbf{T})} \exp \left( -c \sqrt{\frac{2^{2k} \sigma |10 \log \sigma|^2 (\log \log N)^2}{\text{size}(\mathbf{T}')}} \right) |\text{sh}_k(\mathbf{T}')| \\
& \lesssim 2^{2k} \sigma^{10} \sum_{\mathbf{T}' \in \mathcal{T}(\mathbf{T})} e^{-\log \log N} |\text{sh}_k(\mathbf{T}')| \leq 2^{4k} \sigma^{10} |\text{sh}(\mathbf{T})|.
\end{aligned}$$

We used Lemma 6.2, and the fact that  $\text{size}(\mathbf{T}') \leq \sigma$ . Thus, the proof of (7.17), for  $\star = \mathbf{L}$  is complete.

*Proof of the claim.* Ordering  $\mathbf{V} \cap \text{crown}(\mathbf{T}) = \{v_1, v_2, \dots\}$  counterclockwise, we say that  $v_i$  is a cutoff direction if  $\mathbf{T}(v_i) \supsetneq \mathbf{T}(v_{i-1})$ . We show that the sequence  $\{v_i : v_i \text{ is a cutoff}\}$  is lacunary with a top direction of  $\mathbf{T}$ ,  $v_{\mathbf{T}}$ , as its node, and thus it can have at most  $C \log N$  elements, which proves the claim. By pigeonholing, it suffices to prove this under the additional sparseness assumption

$$s, s' \in \mathbf{T}, \omega_{2s} \subset \omega_{2s'} \implies 8|\omega_{2s}| \leq |\omega_{2s'}|.$$

This implies that  $|\text{crown}(\mathbf{T}(v_{i_m}))| \geq 8|\text{crown}(\mathbf{T}(v_{i_{m+1}}))|$  for two consecutive cutoffs  $v_{i_m}$  and  $v_{i_{m+1}}$  with  $i_m < i_{m+1}$ . Thus

$$|v_{i_m} - v_{\mathbf{T}}| \geq \frac{1}{2} |\text{crown}(\mathbf{T}(v_{i_m}))| \geq 4 |\text{crown}(\mathbf{T}(v_{i_{m+1}}))| \geq 2 |v_{i_{m+1}} - v_{\mathbf{T}}|,$$

and this shows that the cutoffs are lacunary with node  $v_{\mathbf{T}}$ . The claim is proved.  $\square$

**7.5. The overlapping part  $\mathbf{O}$ .** We start with the construction of  $\mathcal{F}_{\sigma}^{\mathbf{O}}$ . The trees of  $\mathcal{F}_{\sigma}^{\mathbf{O}}$  will be overlapping trees obtained as saturation of conical trees (see the definitions in Section 6) with conical size roughly  $\sigma$ . Let us give a precise definition of saturation.

Let  $\mathbf{t}$  be a conical tree. Note that  $\{R_s : s \in \mathbf{t}\}$  is a collection of dyadic rectangles with the same shape, coming from a single dyadic grid. Thus, the tiles  $\mathbf{t}_{\max}$  with  $R_s$  maximal are well defined and  $\text{sh}(\mathbf{t})$  is the disjoint union of  $\{R_s : s \in \mathbf{t}_{\max}\}$ . The saturation  $\mathbf{T}(\mathbf{t})$  of a conical tree is defined as

$$\mathbf{T}(\mathbf{t}) = \{s' \in \mathbf{S} : \omega_{1s'} \supset \omega_{\mathbf{t}}, R_{s'} \subset 10R_s \text{ for some } s \in \mathbf{t}_{\max}\}.$$

We omit the dependence on  $\mathbf{t}$  on the notation when no confusion arises. Clearly,  $\mathbf{T}$  is an overlapping tree with top  $\omega_{\mathbf{t}}$  and  $|\text{sh}(\mathbf{T})| \leq 10|\text{sh}(\mathbf{t})|$ .

The next lemma, along the lines of (for example) [31, Lemma 8.2] or [25, Lemma 6.6] will be used in the iterative construction of  $\mathcal{F}_{\sigma}^{\mathbf{O}}$ . We sketch a proof at the end of the subsection.

**Lemma 7.1** (Size Lemma). *Let  $\mathbf{O}' \subset \mathbf{O}$  be a collection of tiles with  $\text{size}_{\nabla}(\mathbf{O}') = \sigma$ . Then  $\mathbf{O}' := \mathbf{O}'_{\text{small}} \cup \mathbf{O}'_{\text{big}}$  where*

$$(7.22) \quad \text{size}_{\nabla}(\mathbf{O}'_{\text{small}}) \leq \frac{\sigma}{2}$$

and  $\mathbf{O}'_{\text{big}} = \cup_{\mathbf{T} \in \mathcal{F}} \mathbf{T}$ . Each  $\mathbf{T} \in \mathcal{F}$  is a saturation of a conical tree  $\mathbf{t}$ , and

$$(7.23) \quad \sum_{\mathbf{T} \in \mathcal{F}} |\text{sh}(\mathbf{T})| \lesssim \sigma^{-2} \|f\|_2^2;$$

Starting with  $\mathbf{O}' = \mathbf{O}$  (which has  $\text{size}_{\nabla}(\mathbf{O}) = 2^K$ ) and iteratively applying Lemma 7.1, we obtain the decomposition  $\mathbf{O} = \cup_{\sigma \leq 2^K} \mathbf{O}_{\sigma}$ , with

$$(7.24) \quad \text{size}(\mathbf{O}_{\sigma}) \leq \sigma,$$

$$(7.25) \quad \mathbf{O}_{\sigma} = \bigcup_{\mathbf{T} \in \mathcal{F}_{\sigma}^{\mathbf{O}}} \mathbf{T}, \text{ and each } \mathbf{T} \in \mathcal{F}_{\sigma} \text{ is the saturation of a conical tree,}$$

$$(7.26) \quad \#(\omega_{\mathbf{T}} := \{\omega : \omega = \omega_{2s} \text{ for some } s \in \tau(\mathbf{t})\}) \leq C \log N,$$

$$(7.27) \quad |\text{sh}(\mathbf{T})| \lesssim \sigma^{-2} \|f\|_2^2.$$

The only property which is not straightforward from the lemma is (7.26). This is an easy consequence of the Vargas set property: the intervals of  $\omega_{\mathbf{T}}$  are pairwise disjoint, and each contains at least one  $v \in \mathbf{V}$ . Denote with  $v_{\ell}$  any of the directions of  $\mathbf{V}$  contained in  $\omega \in \omega_{\mathbf{T}}$ ,  $|\omega| = 2^{-\ell}$ . Then  $2^{-\ell-1} \leq |v_{\ell} - v_{\mathbf{T}}| \leq 2^{-\ell}$ , which means  $\{v_{\ell}\}$  is a lacunary sequence with node  $v_{\mathbf{T}}$ . Thus,  $\{v_{\ell}\}$  (and in turn  $\omega_{\mathbf{T}}$ ) contains at most  $C \log N$  elements.

We now reshuffle the trees in  $\mathcal{F}_{\sigma}^{\mathbf{O}}$  to make sure that

$$(7.28) \quad \sup_{v \in S^1} \sum_{\mathbf{T} \in \mathcal{F}_{\sigma}^{\mathbf{O}}} \mathbf{1}_{\text{crown}(\mathbf{T})}(v) \leq C \log N$$

also holds true. For each dyadic interval  $\omega$ , consider the collection of tiles  $s \in \mathcal{F}_{\sigma}^{\mathbf{O}}$  with  $\omega_{2s} = \omega$ . These tiles are partially ordered by the relation  $s \ll s'$  if  $R_s \subset R'_{s'}$ . Then remove each tile  $s$  which is not maximal with respect to this order relation from the tree  $\mathbf{T}$  to which  $s$  belongs, and assign  $s$  to the tree  $\mathbf{T}'$  where the maximal tile  $s'$  s.t.  $s \ll s'$  belongs. We keep calling the resulting forest  $\mathcal{F}_{\sigma}^{\mathbf{O}}$ . Since this reshuffling does not modify  $\mathbf{O}_{\sigma}$  and since it only makes  $\omega_{\mathbf{T}}$  and  $\text{sh}(\mathbf{T})$  smaller, properties (7.24)-(7.27) still hold.

As a result of the reshuffling, all the tiles  $s \in \mathbf{O}_{\sigma}$  with  $\omega_{2s} = \omega$  will belong to the same tree  $\mathbf{T}$ , and (7.28) follows easily. Suppose that the sum in (7.28) is equal to  $r > 0$ . Then  $v$  belongs to  $r$  (nested) dyadic intervals  $\omega = \{\omega\}$  and for each of them we have  $\omega = \omega_{2s}$  for  $s$  in some (uniquely determined) tree  $\mathbf{T}$ . Then, the collection  $\{\omega_{1s} : \omega_{2s} = \omega, \omega \in \omega\}$  is pairwise disjoint and each  $\omega_{1s}$  contains at least one element of  $\mathbf{V}$ . The same argument used to show (7.26) implies that  $r \leq C \log N$ .

Having at our disposal (7.23) and (7.28), we have established (7.15) and (7.16) for  $* = \mathbf{O}$ . We are left with proving the single tree estimate (7.17) for an overlapping tree  $\mathbf{T} \in \mathcal{F}_{\sigma}$ . We observe that  $\mathbf{T}$  is a disjoint union of (at most  $C \log N$ ) conical trees,

$$\mathbf{T} = \cup_{\omega \in \omega_{\mathbf{T}}} \mathbf{t}(\omega), \quad \mathbf{t}(\omega) := \{s \in \mathbf{T} : \omega_{2s} = \omega\}.$$

The intervals in  $\omega_{\mathbf{T}}$  are pairwise disjoint, so that, with the notation of (6.13),

$$H_{\mathbf{T}(k)}^* f(x) = \sup_{\omega \in \omega_{\mathbf{T}}} |f_{\mathbf{t}(\omega)}^{(k)}(x)|.$$

Then, the estimate

$$|\{x : H_{\mathbf{T}(k)}^* f(x) \gtrsim 2^{2k} \sigma |10 \log \sigma| \log \log N\}| \lesssim \sigma^{10} 2^{2k} |\text{sh}(\mathbf{T})|,$$

which is even stronger than what we need for (7.17), is obtained in the same way as (7.21), applying the estimate for conical trees (6.13) in place of (6.12). This concludes the proof of the theorem.

*Proof of Lemma 7.1.* We insist only on the points which are different from the classical size lemma in [22]; see [31] and [25] for more details.

We can assume that the uncertainty intervals  $\{\omega_{1s} : s \in \mathbf{O}'\}$  come from a single 2-sparse dyadic grid. The following iterative procedure is used to separate  $\mathbf{O}'_{\text{small}}$  from  $\mathbf{O}'_{\text{big}}$ :

- INIT  $\mathcal{F}_+, \mathcal{F}' = \emptyset$ ,  $\text{Stock} := \mathbf{O}'$ ,  $\mathbf{O}'_{\text{big}} := \emptyset$ ;
- WHILE  $\text{size}_{\nabla}(\text{Stock}) \geq \frac{\sigma}{2}$ 
  - select a conical  $\mathbf{t}$  with  $\sum_{s \in \mathbf{t}} |\langle f, \varphi_s \rangle|^2 \geq \frac{\sigma^2}{4} |\text{sh}(\mathbf{t})|$ , with  $\omega_{\mathbf{T}}$  minimal, and  $\text{sh}(\mathbf{T})$  maximal (with respect to inclusion);
  - UPDATE  $\mathcal{F}_+ = \mathcal{F}_+ \cup \mathbf{t}$ ,  $\mathcal{F} = \mathcal{F}' \cup \mathbf{T}(\mathbf{t})$ ,  $\text{Stock} := \text{Stock} - \mathbf{T}(\mathbf{t})$ .

When the algorithm stops, set  $\mathbf{O}'_{\text{small}} := \text{Stock}$ ,  $\mathbf{O}'_{\text{big}} = \cup_{\mathbf{T} \in \mathcal{F}} \mathbf{T}$ ,  $\mathbf{O}'_+ = \cup_{\mathbf{t} \in \mathcal{F}_+} \mathbf{t}$ . The iteration procedure ensures that  $\text{size}_{\nabla}(\mathbf{O}'_{\text{small}}) \leq \frac{\sigma}{2}$ . We are left with proving the inequality

$$(7.29) \quad \sigma^2 \sum_{\mathbf{t} \in \mathcal{F}_+} |\text{sh}(\mathbf{t})| \lesssim \|f\|_2^2.$$

We will use the notation  $F(\mathbf{S}') := \sum_{s \in \mathbf{S}'} \langle f, \varphi_s \rangle \varphi_s$ . Standard arguments reduce the proof of (7.29) to showing that

$$(7.30) \quad \|F(\mathbf{O}'_+)\|_2^2 \lesssim \sigma^2 \sum_{\mathbf{t} \in \mathcal{F}_+} |\text{sh}(\mathbf{t})|.$$

For  $s \in \mathbf{t} \in \mathcal{F}_+$ , define  $\mathbf{B}(s) = \{s' \in \mathbf{O}'_+ - \mathbf{t} : \omega_{1s} = \omega_{1s'}\}$ ,  $\mathbf{C}(s) = \{s' \in \mathbf{O}'_+ - \mathbf{T} : \omega_{1s} \subsetneq \omega_{1s'}\}$ . The tiles  $s' \in \mathbf{B}(s) \cup \mathbf{C}(s)$  are the ones with  $\langle \varphi_s, \varphi_{s'} \rangle \neq 0$ . Then

$$\|F(\mathbf{O}'_+)\|_2^2 \lesssim \sum_{\mathbf{t} \in \mathcal{F}_+} \|F(\mathbf{t})\|_2^2 + \sum_{s \in \mathbf{O}'_+} \langle f, \varphi_s \rangle \langle \varphi_s, F(\mathbf{B}(s)) \rangle + \sum_{s \in \mathbf{O}'_+} \langle f, \varphi_s \rangle \langle \varphi_s, F(\mathbf{C}(s)) \rangle$$

The fact that first and second summand obey the bound of (7.30) is a simple consequence of the almost-orthogonality of the wave packets inside a single conical tree. The third summand is controlled by summing over  $\mathbf{t} \in \mathcal{F}_+$  the estimate

$$\sum_{s \in \mathbf{t}} \langle f, \varphi_s \rangle \langle \varphi_s, F(\mathbf{C}(s)) \rangle \lesssim \sigma^2 |\text{sh}(\mathbf{t})|.$$

Clearly this follows if we establish that, for each  $\bar{s} \in \mathbf{t}_{\text{max}}$ ,

$$(7.31) \quad \sum_{s \in \mathbf{t}(\bar{s})} \langle f, \varphi_s \rangle \langle \varphi_s, F(\mathbf{C}(s)) \rangle \lesssim \sigma^2 |R_{\bar{s}}|.$$

where  $\mathbf{t}(s) := \{s' \in \mathbf{t} : R_s \subset R_{s'}\}$ . Split the collection  $\mathbf{t}(\bar{s})$  into

$$\mathbf{t}(\bar{s}) = \bigcup_{n \geq 0} \mathbf{t}(\bar{s})^{(n)}, \quad \mathbf{t}(\bar{s})^{(n)} := \{s \in \mathbf{t}(\bar{s}) : 2^n R_s \subset 5R_{\bar{s}}, 2^{n+1} R_s \not\subset 10R_{\bar{s}}\}.$$

The rectangles  $\{R_s : s \in \mathbf{t}(\bar{s})\}$  come from a single dyadic grid in the plane, hence

$$\sum_{s \in \mathbf{t}(\bar{s})^{(n)}} |R_s| \lesssim 2^n |R_{\bar{s}}|,$$

so that (7.31) is obtained by summing over  $s \in \mathbf{t}(\bar{s})^{(n)}$  and  $n \geq 0$  the estimate

$$(7.32) \quad \sum_{s' \in \mathbf{C}(s)} |\langle f, \varphi_s \rangle| |\langle \varphi_{s'}, f \rangle| |\langle \varphi_s, \varphi_{s'} \rangle| \lesssim 2^{-10n} \sigma^2 |R_s|.$$

We now show (7.32), beginning with two observations.

OBS.1: let  $s \in \mathbf{t}$  and  $s' \in \mathbf{C}(s)$ . The 2-sparseness property we assumed implies  $\omega_{\mathbf{t}} \subset \omega_{1s'}$ , so that  $\mathbf{t}$  had been selected before the tree containing  $s'$  (by minimality of  $\omega_{\mathbf{t}}$ ). If  $\bar{s} \in \mathbf{t}_{\max}$  is the one tile for which  $s \in \mathbf{t}(\bar{s})$ , the saturation procedure then forces  $8R_{\bar{s}} \cap R_{s'} = \emptyset$ .

OBS.2: by the same token, let  $s, s' \neq s''$  with  $s', s'' \notin \mathbf{t}$ . Then

$$\omega_{1\mathbf{t}} \subset \omega_{1s'} \cap \omega_{1s''}, \text{ ann}(s') = \text{ann}(s'') \implies R_{s'} \cap R_{s''} = \emptyset.$$

Indeed, the 2-sparseness ensures  $\omega_{\mathbf{t}} \subset \omega_{1\mathbf{t}'} \cap \omega_{1\mathbf{t}''}$ . In the case  $\omega_{1\mathbf{t}'} = \omega_{1\mathbf{t}''}$ , the conclusion follows because  $R_{s'}$  and  $R_{s''}$  are dyadic rectangles with the same orientation and side-lengths. If not, assume with no loss in generality that  $\omega_{\mathbf{t}'} \subsetneq \omega_{1\mathbf{t}''}$ , then the conclusion follows exactly as in Observation 1.

Therefore the rectangles  $\{R_{s'} : s' \in \mathbf{C}(s)\}$  are pairwise disjoint (by Observation 2) and do not intersect  $2^n R_s$ , (by Observation 1), which is strictly contained in  $8R_{\bar{s}}$ . Since  $\text{size}_{\nabla}$  controls the size of the single tiles,  $|\langle f, \varphi_s \rangle| \lesssim \sigma \sqrt{|R_s|}$ ,  $|\langle f, \varphi_{s'} \rangle| \lesssim \sigma \sqrt{|R_{s'}|}$ . Using this fact and the rapid decay of the wave packets, (7.32) is a consequence of

$$\sum_{s' \in \mathbf{C}(s)} |R_{s'}| \left(1 + \frac{\text{dist}(R_s, R_{s'})}{|R_s|}\right)^{-100} \lesssim \int_{(2^n R_s)^c} \left(1 + \frac{\text{dist}(x, R_{s'})}{|R_s|}\right)^{-100} dx \lesssim 2^{-10n} |R_s|.$$

This concludes the proof of (7.32), and, in turn, of Lemma 7.1  $\square$

## 8. FINAL REMARKS

Let  $f : \mathbb{R}^2 \rightarrow \mathbb{C}$ . The main unresolved issue in the paper is whether the model sum operator

$$\text{MS}f(x) = \sup_{v \in \mathbf{V}} \left| \sum_{s \in \mathbf{S}_u} \langle f, \varphi_s \rangle \psi_s(x) 1_{\omega_{2s}}(v) \right|$$

has the  $L^2 \rightarrow L^{2,\infty}$  norm of order  $O(\sqrt{\log N} (\log \log N)^{O(1)})$ , for arbitrary  $\mathbf{V} \subset S^1$  with  $N$  elements. Here, as before,  $\varphi_s, \psi_s$  are product wave packets adapted to  $R_s$ .

Consider the related operators

$$\text{S}f(x) = \sup_{v \in \mathbf{V}} \left( \sum_{\text{ann}} \left| \sum_{s \in \mathbf{S}_u : \text{ann}(s) = \text{ann}} \langle f, \varphi_s \rangle \psi_s(x) 1_{\omega_{2s}}(v) \right|^2 \right)^{1/2}$$

$$\text{SQ}f(x) = \sup_{v \in \mathbf{V}} \left( \sum_{s \in \mathbf{S}_u} |\langle f, \varphi_s \rangle|^2 \frac{1_{R_s}(x)}{|R_s|} 1_{\omega_{2s}}(v) \right)^{1/2}$$

$$\mathrm{SC}f(x) = \sup_{v \in \mathbf{V}} \sup_{\omega \subset S^1} \left( \sum_{s \in \mathbf{S}_u: \omega_s = \omega} |\langle f, \varphi_s \rangle|^2 \frac{1_{R_s}(x)}{|R_s|} 1_{\omega_{2s}}(v) \right)^{1/2}.$$

Note the striking simplicity of  $\mathrm{SC}$ . For each  $x$ , the contribution comes from tiles with a fixed eccentricity (i.e. from a single frequency cone).

By invoking a simple randomization argument, it easily follows that

$$\|\mathrm{SC}\|_{L^2 \rightarrow L^{2,\infty}} \lesssim \|\mathrm{SQ}\|_{L^2 \rightarrow L^{2,\infty}} \lesssim \|\mathrm{S}\|_{L^2 \rightarrow L^{2,\infty}} \lesssim \|\mathrm{MS}\|_{L^2 \rightarrow L^{2,\infty}}.$$

The fact that  $\|\mathrm{S}\|_{L^2 \rightarrow L^2} \lesssim \sqrt{\log N}$  was proved in [6], and this is optimal for generic  $\mathbf{V}$ . However, we strongly suspect that  $\|\mathrm{S}\|_{L^2 \rightarrow L^{2,\infty}} \lesssim (\log \log N)^{O(1)}$ . If one proves this, then the conjectured bound  $\|\mathrm{MS}\|_{L^2 \rightarrow L^{2,\infty}} \lesssim \sqrt{\log N} (\log \log N)^{O(1)}$  follows via the Chang-Wilson-Wolff inequality, as described in Section 3. The methods in this paper seem to be unable to conclude even the bound  $\|\mathrm{SC}\|_{L^2 \rightarrow L^{2,\infty}} \lesssim (\log \log N)^{O(1)}$ .

## REFERENCES

- [1] Michael Bateman, *Keakeya sets and directional maximal operators in the plane*, Duke Math. J. **147** (2009), no. 1, 55–77. MR 2494456 (2009m:42029)
- [2] Michael Bateman and Christoph Thiele,  *$L^p$  estimates for the Hilbert transforms along a one-variable vector field*, <http://arxiv.org/abs/1109.6396> (2011).
- [3] S.-Y. A. Chang, J. M. Wilson, and T. H. Wolff, *Some weighted norm inequalities concerning the Schrödinger operators*, Comment. Math. Helv. **60** (1985), no. 2, 217–246. MR 800004 (87d:42027)
- [4] Michael Christ, Javier Duoandikoetxea, and José L. Rubio de Francia, *Maximal operators related to the Radon transform and the Calderón-Zygmund method of rotations*, Duke Math. J. **53** (1986), no. 1, 189–209. MR 835805 (88d:42032)
- [5] A. Cordoba and R. Fefferman, *On the equivalence between the boundedness of certain classes of maximal and multiplier operators in Fourier analysis*, Proc. Nat. Acad. Sci. U.S.A. **74** (1977), no. 2, 423–425. MR 0433117 (55 #6096)
- [6] Ciprian Demeter, *Singular integrals along  $N$  directions in  $\mathbb{R}^2$* , Proc. Amer. Math. Soc. **138** (2010), no. 12, 4433–4442. MR 2680067 (2011i:42023)
- [7] ———,  *$L^2$  bounds for a Keakeya type maximal operator in  $\mathbb{R}^3$* , <http://arxiv.org/abs/1105.1115v1> (2011).
- [8] Ciprian Demeter, Michael T. Lacey, Terence Tao, and Christoph Thiele, *Breaking the duality in the return times theorem*, Duke Math. J. **143** (2008), no. 2, 281–355. MR 2420509 (2009f:42013)
- [9] Javier Duoandikoetxea and Ana Vargas, *Directional operators and radial functions on the plane*, Ark. Mat. **33** (1995), no. 2, 281–291. MR 1373025 (97c:42031)
- [10] R. Fefferman and J. Pipher, *Multiparameter operators and sharp weighted inequalities*, Amer. J. Math. **119** (1997), no. 2, 337–369. MR 1439553 (98b:42027)
- [11] José García-Cuerva and José L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Mathematics Studies, vol. 116, North-Holland Publishing Co., Amsterdam, 1985, Notas de Matemática [Mathematical Notes], 104. MR 807149 (87d:42023)
- [12] Loukas Grafakos, Petr Honzík, and Andreas Seeger, *On maximal functions for Mikhlin-Hörmander multipliers*, Adv. Math. **204** (2006), no. 2, 363–378. MR 2249617 (2007d:42015)
- [13] Richard A. Hunt, *An estimate of the conjugate function*, Studia Math. **44** (1972), 371–377, Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, IV. MR 0338667 (49 #3431)
- [14] T. Hytonen, M. Lacey, H. Martikainen, T. Orponen, M. Reguera, and E. Sawyer, *Weak and strong type estimates for maximal truncations of Calderón-Zygmund operators on  $A_p$  weighted spaces*, <http://arxiv.org/abs/1103.5229v1> (2011).
- [15] G. A. Karagulyan, *On unboundedness of maximal operators for directional Hilbert transforms*, Proc. Amer. Math. Soc. **135** (2007), no. 10, 3133–3141 (electronic). MR 2322743 (2008e:42044)

- [16] B. S. Kashin and A. A. Saakyan, *Orthogonal series*, Translations of Mathematical Monographs, vol. 75, American Mathematical Society, Providence, RI, 1989, Translated from the Russian by Ralph P. Boas, Translation edited by Ben Silver. MR 1007141 (90g:42001)
- [17] Nets Hawk Katz, *Maximal operators over arbitrary sets of directions*, Duke Math. J. **97** (1999), no. 1, 67–79. MR 1681088 (2000a:42036)
- [18] ———, *Remarks on maximal operators over arbitrary sets of directions*, Bull. London Math. Soc. **31** (1999), no. 6, 700–710. MR 1711029 (2001g:42041)
- [19] Joonil Kim, *Sharp  $L^2$  bound of maximal Hilbert transforms over arbitrary sets of directions*, J. Math. Anal. Appl. **335** (2007), no. 1, 56–63. MR 2340304 (2009a:42012)
- [20] Michael Lacey and Xiaochun Li, *On a conjecture of E. M. Stein on the Hilbert transform on vector fields*, Mem. Amer. Math. Soc. **205** (2010), no. 965, viii+72. MR 2654385 (2011c:42019)
- [21] Michael Lacey and Jason Metcalfe, *Paraproducts in one and several parameters*, Forum Math. **19** (2007), no. 2, 325–351. MR 2313844 (2008b:42026)
- [22] Michael Lacey and Christoph Thiele, *A proof of boundedness of the Carleson operator*, Math. Res. Lett. **7** (2000), no. 4, 361–370. MR 1783613 (2001m:42009)
- [23] Michael T. Lacey and Xiaochun Li, *Maximal theorems for the directional Hilbert transform on the plane*, Trans. Amer. Math. Soc. **358** (2006), no. 9, 4099–4117 (electronic). MR 2219012 (2006k:42018)
- [24] Camil Muscalu, Jill Pipher, Terence Tao, and Christoph Thiele, *Multi-parameter paraproducts*, Rev. Mat. Iberoam. **22** (2006), no. 3, 963–976. MR 2320408 (2008b:42037)
- [25] Camil Muscalu, Terence Tao, and Christoph Thiele,  *$L^p$  estimates for the biest. II. The Fourier case*, Math. Ann. **329** (2004), no. 3, 427–461. MR 2127985 (2005k:42054)
- [26] S. Petermichl, *The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical  $A_p$  characteristic*, Amer. J. Math. **129** (2007), no. 5, 1355–1375. MR 2354322 (2008k:42066)
- [27] Jill Pipher, *Bounded double square functions*, Ann. Inst. Fourier (Grenoble) **36** (1986), no. 2, 69–82. MR 850744 (88h:42021)
- [28] Elias M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. MR 0290095 (44 #7280)
- [29] ———, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. MR 1232192 (95c:42002)
- [30] Elias M. Stein and Brian Street, *Multi-parameter singular Radon transforms*, Math. Res. Lett. **18** (2011), no. 2, 257–277. MR 2784671 (2012b:44007)
- [31] Christoph Thiele, *Wave packet analysis*, CBMS Regional Conference Series in Mathematics, vol. 105, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006. MR 2199086 (2006m:42073)

DEPT. OF MATHEMATICS  
 INDIANA UNIVERSITY  
 BLOOMINGTON, IN 47405 - USA  
*E-mail address:* demeter@indiana.edu (C. Demeter)

DEPT. OF MATHEMATICS & INSTITUTE FOR SCIENTIFIC COMPUTING AND APPLIED MATHEMATICS  
 INDIANA UNIVERSITY  
 BLOOMINGTON, IN 47405 - USA  
*E-mail address:* fradipli@indiana.edu (F. Di Plinio)